

BIROn - Birkbeck Institutional Research Online

Zivony, Alon and Eimer, Martin (2021) Distractor intrusions are the result of delayed attentional engagement: a new temporal variability account of attentional selectivity in dynamic visual tasks. *Journal of Experimental Psychology: General* 150 (1), pp. 23-41. ISSN 0096-3445.

Downloaded from: <https://eprints.bbk.ac.uk/id/eprint/31560/>

Usage Guidelines:

Please refer to usage guidelines at <https://eprints.bbk.ac.uk/policies.html>
contact lib-eprints@bbk.ac.uk.

or alternatively

Running Head: DISTRACTOR INTRUSIONS AND ATTENTIONAL ENGAGEMENT

Distractor intrusions are the result of delayed attentional engagement:
A new temporal variability account of attentional selectivity in dynamic visual tasks

Alon Zivony* & Martin Eimer

Department of Psychological Sciences, Birkbeck College, University of London,
Malet Street, London WC1E 7HX, UK

*This is the accepted authors' manuscript, not the final published article. The published article is
available at:*

Journal of Experimental Psychology: General

Word count: 12,582

* Corresponding author

Phone: 0044 20 76316522

Fax: 0044 20 76316312

Email: alonzivony@gmail.com

Abstract

When observers have to identify targets among distractors in a rapid serial visual presentation (RSVP) stream, distractor intrusion errors are frequent, demonstrating the difficulty of allocating attention to the right object at the right moment in time. However, the mechanisms responsible for such intrusion errors remain disputed. We propose a new attentional engagement account of selective visual processing in RSVP tasks. Engagement is triggered by the pre-attentive detection of target-defining features. Critically, the success versus failure of target identification is determined by the speed of such engagement processes on individual trials. To test this account, we measured electrophysiological markers of attentional engagement (N2pc components) in three experiments where observers had to report the identity of a target digit in one of two lateral RSVP streams. On most trials, the target was immediately followed by a digit distractor, resulting in many post-target distractor intrusions. Critically, N2pc components measured on distractor intrusion trials were significantly delayed relative to trials with correct target reports. This was the case regardless of whether the target was defined by a shape cue or by its colour, and even when the location of shape-defined targets was known in advance. These findings show that distractor intrusions are the result of delayed attentional engagement. They demonstrate that temporal variability in attentional selectivity across trials can strongly affect visual awareness and perceptual reports. Our temporal variability account of attentional engagement offers a new framework for assessing the temporal dynamics of attention in visual object recognition.

Keywords: temporal attention, N2pc, distractor intrusion, RSVP, attentional engagement

The function of selective visual attention is to enable the detection and identification of task-relevant visual objects, and the filtering of other objects that are not relevant to current task goals. Selective attention is particularly important when multiple competing objects are present at the same time, or when these objects appear sequentially and in rapid succession. The former situation has been extensively studied in visual search tasks, where a target object is presented simultaneously with multiple distractors. In such tasks, attention can be directed rapidly to targets with a distinctive attribute, but attentional guidance is less effective when targets and distractors share one or several features (e.g., Duncan & Humphreys, 1989; Treisman & Gelade, 1980; Wolfe, 2014). The main challenge for selective attention in visual search is spatial uncertainty, as the location of a target object in a particular search display is not known in advance. A different problem arises in situations where multiple objects appear and disappear in rapid succession, and one of these objects has to be identified. In such rapid serial visual presentation (RSVP) tasks, temporal uncertainty is the main challenge for attentional selectivity, as the position of target objects within an RSVP stream is usually not predictable. To identify these targets, attention has to be allocated to the right object at the right moment in time.

Models of temporal attentional selection processes emphasize the fact that a single target can be easily differentiated from distractors in streams where items appear at a frequency of approximately 10 items per second (Chun & Potter, 1995; Di Lollo, Kawahara, Shahab Ghorashi & Enns, 2005; Jolicoeur & Dell'Acqua, 1998; Taatgen, Juvina, Schipper, Borst, & Martens, 2009). However, studies that demonstrate such apparently efficient temporal selection have employed distractors that are categorically distinct from the target and therefore do not share its response-relevant dimension (e.g., a target digit that is embedded among letters). In contrast, the ability to select a target and ignore temporally adjacent distractors is substantially reduced under

conditions where targets and distractors share the same task-relevant category, and the identity of a distractor can therefore be reported (e.g., a target digit embedded among distractor digits). In such tasks, participants will often erroneously report the identity of temporally adjacent distractors. Despite the robustness of this distractor intrusion phenomenon (e.g., Botella & Eriksen, 1992; Botella, Barriopedro & Suero, 2001; Chun, 1997; Goodbourn & Holcombe, 2015; Intraub, 1985; Gathercole & Broadbent, 1984; Popple & Levi, 2007; Recht, Mamassian & de Gardelle, 2019; Vul, Hanus & Kanwisher, 2009), its theoretical implications for models of how selective attention operates in the time domain have been largely neglected.

One reason why the general significance of distractor intrusions has not been recognized is that they could in principle be regarded as a result of rare failures of temporal selectivity. Most studies of distractor intrusions do not directly investigate whether and how potentially intruding distractors disrupt the efficiency of temporal target selection processes. They focus instead on the frequency of distractor reports as a function of their temporal lag from the target, and measure positional errors in order to determine the distribution of temporal attention across multiple items. Such errors can indicate whether attention is temporally centred on the target (e.g., Chun, 1997; Goodbourn & Holcombe, 2015) or allocated more broadly across multiple successive items (e.g., Vul, Nieuwenstein & Kanwisher, 2008). However, they do not provide any direct insights into the causes of distractor intrusion errors, and whether such errors reflect general and theoretically important limitations of temporal selection.

Another reason why the relevance of distractor intrusion errors for models of attentional selectivity in the time domain has not been acknowledged is that there is as yet no generally accepted account of the mechanisms responsible for such errors. One explanation for distractor intrusions was proposed by Botella and colleagues (Botella et al., 2001; Botella, Arend, & Suero,

2004). According to these authors, the presence of a task-relevant selection feature is detected by pre-attentive visual processes, and this results in an attempt to focus attention on the target. When this attempt is successful, the target is correctly identified. On a subset of trials, attentional focusing fails, and perceptual reports have to be based on a sophisticated guessing mechanism. This mechanism selects one of the currently available visual representations of items in the RSVP stream, resulting in intrusion errors when a distractor representation is picked. A similar account was proposed by Vul et al. (2009). They postulate that once a selection feature is detected, several item representations within a time window centred on the target are selected and stored in short-term memory. The strength of the representations within this selection window varies in a graded fashion, with the target object usually most strongly activated. One of these representations gains access to subsequent identification and report processes, and this is determined by a probabilistic sampling process. In the majority of trials, the most strongly activated representation (i.e., the target) is sampled, but occasionally one of the distractor representations is picked instead, producing intrusion errors.

It is notable that neither of these two accounts of distractor intrusion errors refer specifically to temporal aspects of attentional processing. Botella et al. (2001) attribute distractor intrusions to a general failure of attentional focusing in some trials. According to Vul et al. (2009), intrusion errors are produced by a post-perceptual sampling process, and are unrelated to the speed of prior attentional selection mechanisms. These authors assume that the temporal position of the attentional selection window that provides the input to the subsequent sampling process does not vary from trial to trial. They based this conclusion on an experiment where observers had to provide their best guess about the identity of the target in an RSVP stream, followed by an additional second guess. If the temporal position of the selection window varied considerably

across trials, the first report should predict the temporal position of the second item that was reported. No such temporal associations between the two reports were found, suggesting little trial-by-trial variation in the position of the attentional window.

Though it is undeniable that perceptual processes vary in their efficiency on trial-by-trial basis, most models view such variability as no more than a source of statistical error (Ashby & Townsend, 1986; Ashby & Lee, 1993). However, it is possible that variability in the speed of attentional selection plays an important role in determining the content of conscious perception (e.g., Hogendoorn, Carlson, & Verstraten, 2011). Given that the main challenge for attention in RSVP tasks is to select the correct object at the right moment in time, the presence of distractor intrusions might be systematically linked to differences in the timing of attentional selection processes across individual trials. The goal of the present study was to provide evidence for such an alternative account.

In line with previous suggestions (e.g., Chun & Potter, 1995; Wyble, Bowman & Nieuwenstein, 2009), we assume that pre-attentive visual processing produces short-lived representations of individual items within an RSVP stream. Once a task-relevant selection feature (e.g., a specific colour) is detected, a transient attentional facilitation of visual activity rapidly builds up, enhances the strength of the available stimulus representations, and then gradually dissipates (Reeves & Sperling, 1986; Shih & Sperling, 2002; Sperling & Weichselgartner, 1995; Weichselgartner & Sperling, 1987, see also: Wyble et al., 2009; Wyble, Potter, Bowman & Nieuwenstein, 2011). We use the term *attentional engagement* to describe the start of this attentional amplification (e.g., Posner & Peterson, 1990; Folk, Ester & Troemel, 2009; Nieuwenstein, Chun, van der Lubbe & Hooge, 2005; Zivony & Lamy, 2016; 2018), and the term *attentional episode* to refer to the interval between the onset and offset of this response

(Wyble et al., 2011)¹. Distractor intrusions arise on trials where attentional engagement processes fail to make the response features of a target object accessible to subsequent identification and verbal report mechanisms, and a distractor object is reported instead. Our critical new hypothesis is that fluctuations in the onset of attentional engagement across individual trials determine whether a target is identified correctly or a distractor intrusion occurs. Attentional engagement is triggered once sufficient perceptual evidence for the presence of a selection feature has been accumulated (Zivony & Lamy, 2018), and this point in time varies across trials. In RSVP tasks where target and distractor objects are presented in rapid succession, this temporal variability in attentional engagement will affect the relative strength of target and distractor representations and thus the accuracy of perceptual reports. On trials where attentional engagement is fast, correct target reports are very likely. In contrast, on trials with slower engagement, the probability of reporting the distractor that follows the target increases.

We conducted three experiments that tested this new temporal variability account with on-line electrophysiological markers of attentional engagement. These markers were obtained separately for trials with correct target reports and trials with post-target distractor intrusions, in order to assess whether these two types of trials differ systematically with respect to the speed of attentional engagement processes. In Experiment 1, participants had to report the identity of a target digit that was embedded in one of two concurrent RSVP streams in the left and right visual field. The target was surrounded by a pre-specified outline shape (circle or square), which served as the selection feature (see Figure 1). Prior to the target's appearance, most distractors were

¹ The concepts of an attentional episode and attentional engagement need to be distinguished because the onset of an attentional episode can be delayed relative to the onset of attentional engagement if the attentional response is disrupted by external factors (Zivony & Lamy, 2016), and because objects that appear after a target can extend the duration of an attentional episode without necessarily triggering attentional engagement (Callahan-Flintoft, Chen & Wyble, 2018; Tan & Wyble, 2015).

letters and thus did not share the target's response feature. On 25% of all trials, the target was followed by a letter, which allowed us to estimate baseline accuracy levels on trials without any competition between the target and temporally adjacent distractors. However, on 75% of all trials, the object that immediately followed the target in the same stream was also a digit. Here, post-target intrusion errors were possible, where participants report the identity of this distractor instead of the preceding target digit.

To track the speed of attentional engagement processes, we recorded EEG during task performance, and measured N2pc components by comparing ERPs at posterior electrodes contralateral and ipsilateral to the visual field where the target appeared. The N2pc is an established electrophysiological marker of the allocation of attention to visual objects with task-relevant features (e.g., Eimer, 1996, Woodman & Luck, 1999), and has previously also been employed to track the time course of attentional engagement (e.g., Callahan-Flintoft, Chen & Wyble, 2017; Zivony, Allon, Luria & Lamy, 2018; see also Kiss, Van Velzen & Eimer, 2008). Given these previous findings, an N2pc component should be elicited by the RSVP frame that contains the selection feature and the target object. Importantly, the onset latency of this N2pc marks the point in time where attentional engagement processes are activated. To test whether trial-by-trial variability in the onset of these engagement processes determines the occurrence of post-target intrusions, N2pcs were measured separately for trials where the identity of the target was correctly reported, and for trials where the identity of the post-target distractor digit was reported instead. We focused exclusively on post-target distractor intrusions in this study, because previous experiments have shown they are the most common type of intrusion errors (e.g., Botella et al., 2001; Goodbourn & Holcombe, 2015; Recht et al., 2019). Given the frequency of these errors, they should occur on a sufficient number of trials to allow for

calculating meaningful N2pc components for each participant, in spite of the fact that error rates are likely to vary considerably across participants.

If distractor intrusions reflected a complete failure of targets to engage attention (Botella et al., 2001), no N2pc should be observed at all on distractor intrusion trials. If distractor intrusions were entirely unrelated to the temporal variability of attentional engagement across trials (as proposed by Vul et al., 2009), trials with correct responses and distractor intrusions should not differ in terms of target N2pc latencies. In contrast, the hypothesis that post-target intrusion errors occur when attentional engagement is delayed predicts systematic differences in N2pc onset latencies between these two types of trials. N2pc components should emerge significantly later on distractor intrusion trials relative to trials with correct responses.

Experiment 1

Because this is the first study that examined possible N2pc differences between trials with correct responses and distractor intrusion errors, we could not conduct a power analysis based on previous N2pc results from similar experiments to justify our sample size. Therefore, Experiment 1 was divided to two parts. In Experiment 1A, we conducted an exploratory study with a sample size of $N = 12$ (which is in line with previous studies from our lab that examine within-subject modulations of the N2pc component, e.g., Berggren & Eimer, 2019; Jenkins, Grubert, & Eimer, 2018). The results of this study were then used to determine the appropriate sample size for the following experiments, including Experiment 1B, which was a direct replication of Experiment 1A.

Experiment 1A: Method

Participants

Participants were 12 (8 women) volunteers ($M_{age} = 28.92$, $SD = 8.54$) who participated for £25. All reported normal or corrected-to-normal visual acuity.

Apparatus

Stimuli were presented on a 24-inch BenQ LED monitor (100 Hz; 1920 x 1080 screen resolution) attached to a SilverStone PC, with participant viewing distance at approximately 80 cm. Manual responses were registered via a standard computer keyboard.

Stimuli and design

All methods used in this experiment, and subsequent experiments, were approved by the institution's departmental ethical guidelines committee at Birkbeck, University of London. Participants had to report as accurately as possible the numerical value of a digit (report feature) that appeared inside a pre-specified shape (circle or square; selection feature), by pressing the corresponding keyboard button. These targets were presented unpredictably in one of two RSVP streams on the left and right side. Manual responses were executed without time pressure at the end of each trial. The sequence of events is illustrated in Figure 1. Each trial began with the presentation of a fixation display (a grey $0.2^\circ \times 0.2^\circ$ "+" sign at the center of the screen). Then, after 500 ms, two lateral RSVP streams including 8 to 11 frames appeared along with the fixation cross. Each frame appeared for 50 ms, followed by an ISI of 50 ms. The response screen was identical to the fixation display and remained present until a response was registered. Following this response, a blank screen appeared for 800 ms before a new trial started.

All stimuli in the RSVP streams were grey (CIE colour coordinates: 0.309/.332, luminance

46.6 cd/m²). Each frame consisted of two alphanumeric characters (1.3° in height) appearing at a center-to-center distance of 4.5° to the left and right of fixation. Letters in each stream were randomly selected without replacement from a 23-letter set (all English alphabet letters, excluding I, X, and O), with the sole restriction that the same letter could not appear in both streams at the same time. Digits were selected without replacement from a set of 6 digits (2, 3, 4, 6, 7, and 8). The target digit appeared with equal probability and unpredictably in the 5th, 6th, 7th, or 8th frame within the RSVP stream, either in the left or right RSVP stream. This target frame contained one digit and one letter, which appeared within two different outline shapes (square: 1.5° in side, and circle: 1.68° in diameter, line width for both: 4 pixel). The digit was always presented within the pre-specified target shape, and the letter within the other shape. The frame immediately preceding the target frame always included two letters (to prevent any pre-target intrusion errors). The earlier pre-target frames were equally likely to contain two letters, or one digit and one letter (with digit and letter location randomly selected for each frame). The target frame was always followed by two additional frames. On 75% of all trials, the frame immediately following the target contained a digit in the same location as the preceding target digit, so that post-target distractor intrusion errors were possible (Figure 1A). On the remaining 25% randomly intermixed trials, this frame contained two letters (Figure 1B). The next two and final frames always included two letters.

The experiment included 10 practice trials followed by 600 experimental trials, divided into 50-trial blocks. For half the participants, the target-defining selection feature was the square for the first 6 blocks and the circle for the rest. For the other half, this order was reversed. Instructions about this shape change were given before the beginning of the 7th block, followed by additional 5 practice trials. Participants were allowed to take self-paced breaks between

blocks. They were informed that target digits were equally likely to appear in the left or right RSVP stream, and that task-irrelevant digits would appear prior to the target. This ensured that attentional allocation processes would be guided by the selection feature (circle or square), rather than by alphanumeric category (i.e., attending to the first digit in the stream).

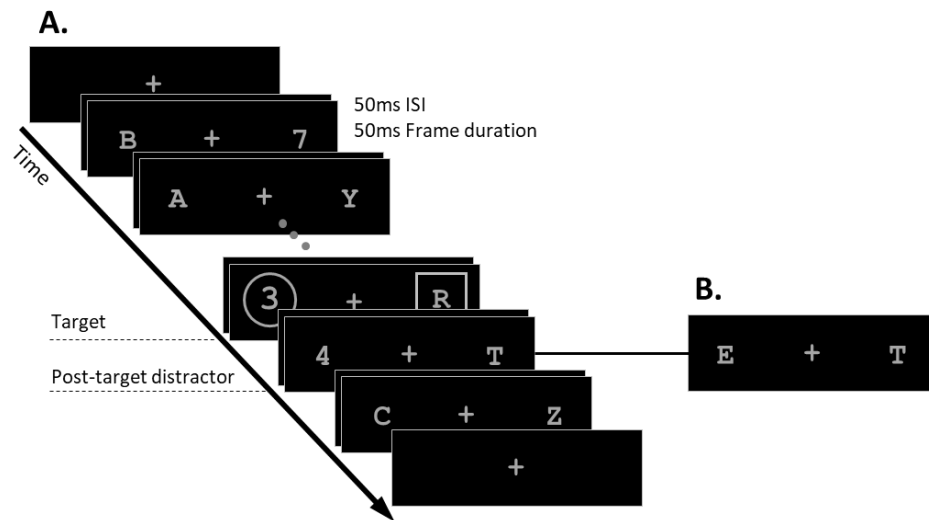


Figure 1. Illustration of the stimulus sequence in Experiment 1 and 3. Participants had to report the target digit within one of two RSVP streams, defined by a pre-defined selection feature (e.g., circle). The target appeared at positions 5 to 8 within the stream, and was followed by two additional frames. The post-target frame contained a digit at the same location as the target on 75% of trials (A) and two letters on 25% of trials (B).

EEG Recording and Data Analysis

EEG was DC-recorded from 27 scalp electrodes, mounted on an elastic cap at sites Fpz, F7, F8, F3, F4, Fz, FC5, FC6, T7, T8, C3, C4, Cz, CP5, CP6, P9, P10, P7, P8, P3, P4, Pz, PO7, PO8, PO9, PO10, and Oz. A 500-Hz sampling rate with a 40 Hz low-pass filter was applied. Channels were referenced online to a left-earlobe electrode, and re-referenced offline to an average of both earlobes. No other filters were applied after EEG acquisition. Trials with eye blinks (exceeding

$\pm 60 \mu\text{V}$ at Fpz), horizontal eye movements (exceeding $\pm 30 \mu\text{V}$ in the HEOG channels), and muscle movement artifacts (exceeding $\pm 80 \mu\text{V}$ at all other channels) were removed as artifacts. EEG was segmented into epochs from 100 ms before to 500 ms after the onset of the target frame, relative to a 100 ms pre-stimulus baseline. ERPs for the majority of trials where the post-target distractor was a digit were computed separately for trials where the target digit was reported correctly and for trials where the identity of the post-target distractor digit was reported instead. Trials where participants reported neither the target nor this distractor were excluded. Averaged ERP waveforms were computed for trials with a target in the left or right RSVP stream, and N2pc components triggered by the target frame were computed by comparing ERPs at electrodes PO7/PO8 contralateral and ipsilateral to the location of the target.

N2pc analyses. Analyses focused on trials with post-target digit distractors, and compared N2pcs on trials with correct responses versus intrusion errors. N2pc onset latencies on these trials were calculated on the basis of contralateral-ipsilateral difference waveforms, following (i) an application of a 10-hz low pass filter (Brisson, Robitaille & Jolicœur, 2007)² to all electrodes other than the HEOG electrodes prior to segmentation, and (ii) the jackknife-based procedure described by Miller, Patterson, and Ulrich (1998). We defined the N2pc onset criterion as the point where the difference waveform reached 50% of the average N2pc peak amplitude across trials with correct responses and distractor intrusion trials, which provides an estimate of the average onset time across trials (Luck, 2014). There were three reasons for using the average peak across correct and intrusion trials for our calculation, instead of the peak for each condition. Using the same criteria for both conditions avoids a distortion due to differences in the N2pc amplitude (see Grubert & Eimer, 2015; Grubert, Krummenacher & Eimer, 2011). Unlike

² All N2pc latency differences between trials with correct responses versus distractor intrusions reported in this study were replicated when no additional filter was applied.

fractional area analyses, this method is insensitive to selected time windows and to any negativity that emerges after the N2pc (as long as the target-locked N2pc is the largest negative component). Relative to a constant criteria (e.g., using an a-priori criteria of $-1 \mu\text{V}$ for all conditions), this method allows for a better comparison between conditions that yield peak N2pc amplitudes of variable sizes. Since the jackknife procedure greatly reduces error variance, in all statistical N2pc onset analyses, F scores were adjusted according to the formula provided by Ulrich and Miller (2001). N2pc amplitudes were defined at the mean amplitude of the ipsilateral-contralateral difference waveform in the 200–300 ms time window after the onset of the target frame. Both the choice of electrode sites and time-window for the N2pc are standard in our lab (e.g., Berggren & Eimer, 2019; Kiss et al., 2008) and are frequently used in other studies of the N2pc (e.g., Callahan-Flintoft et al., 2017; see also: Luck, 2014).

Residual eye movement analysis. While our exclusion criteria for eye movements ensured that no large movement affected our results, it is possible that small but consistent eye movements in the direction of a target may have been left in the data (Lins, Picton, Berg & Scherg, 1993). To ensure that these small eye movements did not create any systematic N2pc differences between correct trials and intrusion trials, we analysed data from the two HEOG electrodes ipsilateral and contralateral to the visual field where the target appeared. We calculated the difference wave between the ipsilateral and contralateral HEOG traces, such that a positive deflection indicates a tendency for a small deviation of eye gaze towards the target. We then examined whether averaged HEOG difference waves differed between correct and intrusion trials during the N2pc time window. This analysis, reported in the Supplementary materials, suggested that any residual small eye gaze deviations remaining in the data did not contribute to the N2pc differences between the two conditions in any of the experiments reported here.

Statistical analysis of null results

Since the absence of a significant effect does not constitute as evidence in favour of the null hypothesis, statistical tests with non-significant results were supplemented, when possible, with a corresponding calculation of a Bayes Factor in favour of the null hypothesis (BF_{01}). All tests were conducted using JASP (0.9.2). Differences between two groups were tested with a dependent-sample Bayesian t-test. Bayes Factors associated with a two-way interaction were calculated by dividing two Bayes Factors: (i) the Bayes Factor associated with the full model (including the interaction and both main effects), and (ii) the Bayes Factor associated with the model that includes only the two main effects (Wagenmakers et al., 2018). Bayes Factors associated with a main effect in a two-way design were isolated by dividing the model with both main effects and the model with the irrelevant main effect. Since we had no a-priori expectations regarding these effects, we used default priors for all of these tests (Cauchy scale of 0.707 for t-tests and $r_A = 0.5$ for ANOVAs).

Experiment 1A: Results*Behavioural results*

Preliminary analysis indicated that the shape of the selection feature (square vs. circle) had no effect on accuracy rates, $F < 1$, $BF_{01} = 4.685$, and did not interact with distractor type (letter vs. digit), $F < 1$, $BF_{01} = 3.79$. Therefore all data were collapsed across this factor. As expected, response accuracy was impaired when the target was followed by a digit distractor relative to when it was followed by a letter distractor (percentage correct: $M = 36.1\%$ vs. $M = 77.2\%$, $t(11) = 11.39$, $p < .001$). When the post-target distractor was a digit, 57.0% of the responses were distractor intrusions accounting for 89.2% of all errors on these trials). Mean accuracy and

intrusion rates are presented in Figure 2A (see Supplementary Figure 2 for individual results).



Figure 2. Frequency of correct responses and distractor intrusions in all three experiments, as a function of the post-target distractor identity (letter vs. digit). For Experiment 1, results from Experiment 1A and 1B are shown separately. For Experiment 2, results from blocks with grey versus coloured post-target distractors are shown separately. Error bars denote one standard error.

N2pc components

The average general EEG data loss due to artifacts was 10.8% ($SD = 11.3\%$). Figure 3 (left panels) shows the ERP waveforms triggered by the target frame at electrodes PO7 and PO8 contralateral and ipsilateral to the target, for the 75% of all trials where the target digit was followed by a distractor digit at the same location. ERPs are presented separately for trials with correct responses and distractor intrusion errors. The corresponding difference waves obtained by subtracting ipsilateral from contralateral ERPs are shown in Figure 4A (see Supplementary Figure 2 for individual waveforms). Clear N2pc components were present for both types of trials, but there was a marked N2pc onset latency difference, with an N2pc delay on trials where distractor intrusions were reported. This was confirmed by the analysis of N2pc onset latencies, based on a 50% average peak amplitude criterion ($M = -1.04 \mu V$). The N2pc component emerged approximately 30 ms earlier on trials with correct responses relative to distractor intrusion trials,

$M = 214.9$ ms vs. $M = 244.75$ ms, and this difference was significant, $F_{adjusted}(1,11) = 13.93$, $p = .003$. N2pc mean amplitudes measured in the 200–300 ms time window were significantly different from zero both on trials with correct responses and on distractor intrusion trials, $p < .001$ and $p = .001$, respectively. However, N2pcs were reliably larger on trials where the target was reported correctly, $t(11) = 4.31$, $p = .001$.

EXPERIMENT 1

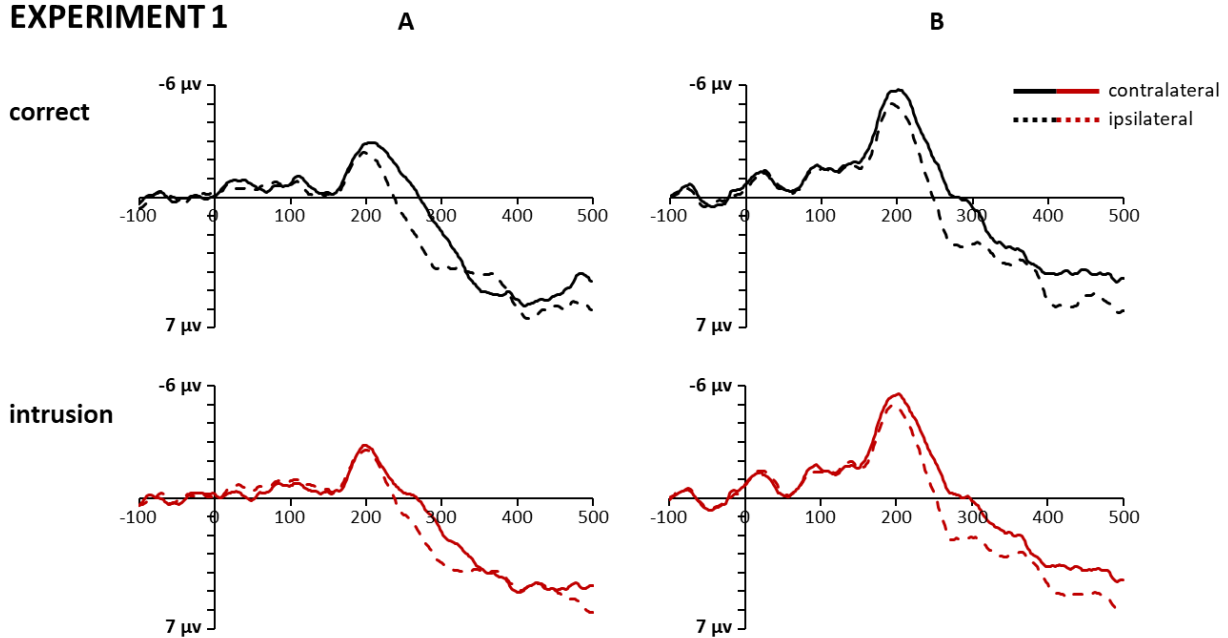


Figure 3. Grand-average event-related potentials (ERPs) waveforms elicited in Experiment 1 by target frames at electrodes PO7/PO8 on contralateral and ipsilateral electrodes relative to the target. ERPs for RSVP streams with a post-target digit distractor are shown separately for Experiment 1A (left panels) and Experiment 1B (right panels) and separately for trials with correct responses (top panels) and distractor intrusion trials (lower panels).

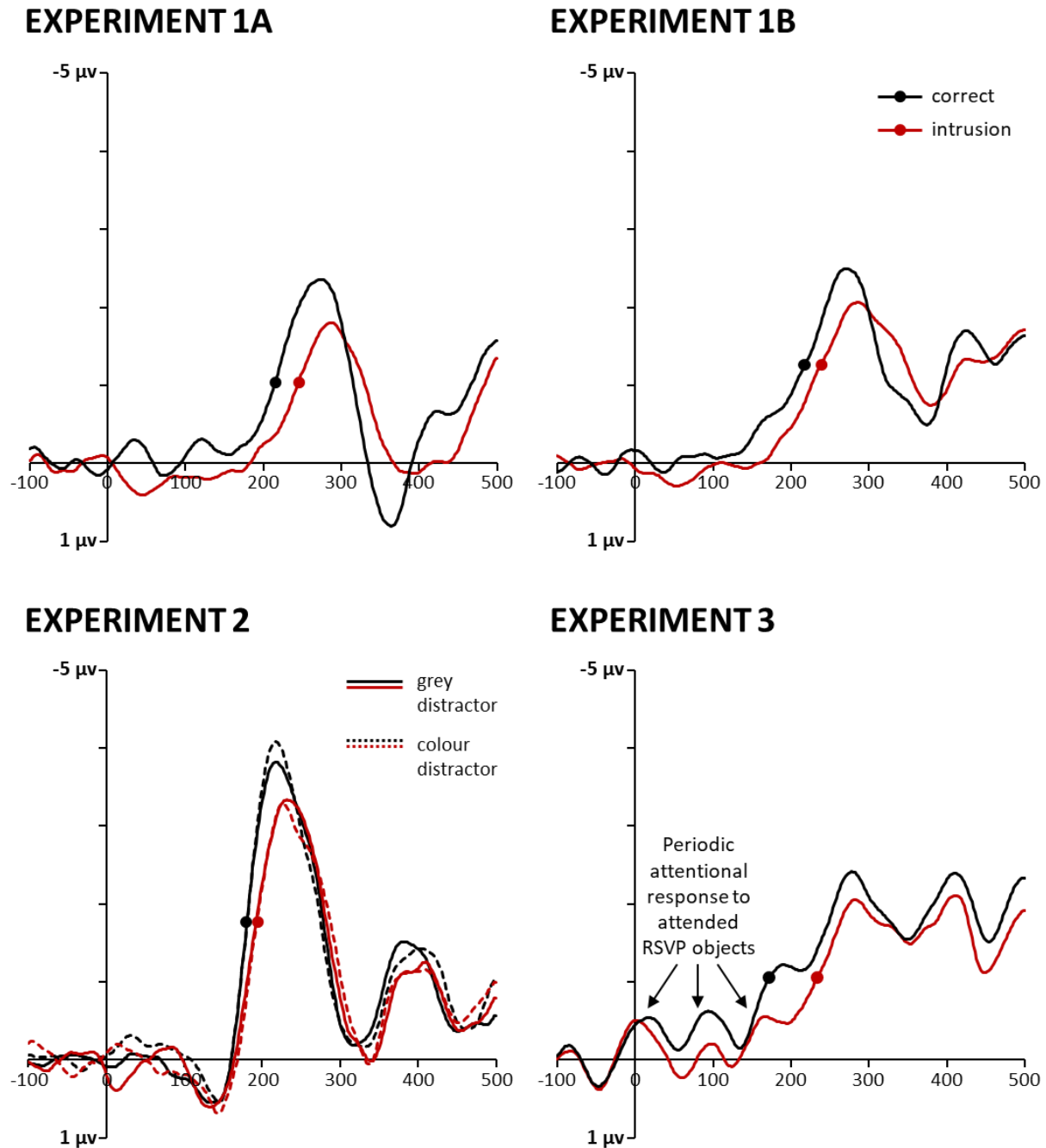


Figure 4. N2pc difference waveforms obtained by subtracting ipsilateral from contralateral ERPs measured in response to RSVP streams with a post-target digit distractor. Waveforms are time-locked to the onset of the target frame, and are shown separately for trials with correct responses and distractor intrusion trials, for all three experiments. N2pc onset latencies are indicated by dots. In line with the N2pc onset analyses, a 10 Hz low-pass filter was applied to these waveforms.

Experiment 1B: Method

Sample size selection

On the basis of Experiment 1A, we calculated the sample size required to observe significant differences in target-locked N2pc onset latency between trials with correct responses and distractor intrusions. We conducted this analysis with G*Power (Faul et al., 2013), using an alpha of .05 and power of .80. Because the onset latency analysis in Experiment 1A was based on jackknifed N2pc waveforms, it is questionable whether the effect size (as reflected by η_p^2) is meaningful in any context other than determining sample size for a similar analysis. Based on the results reported above [i.e., $F_{adjusted}(1,11) = 13.93$], this effect size was calculated to be $\eta_p^2 = .558$. Based on these data, the power analysis for a repeated measures F-test yielded a minimum required sample size of 10 participants. For comparability with Experiment 1A, we decided to again test a sample of 12 participants, which yielded actual power of 92%.

Participants

Participants were 12 (8 women) volunteers ($M_{age} = 30.33$, $SD = 10.90$) who participated for £25. All reported normal or corrected-to-normal visual acuity. One participant was excluded from analysis because their low accuracy rate on trials where the post-target distractor was a digit (13.5%; 69.4% intrusions) and their high rejection rate due to eye blinks and eye movements (54.1%) left too few trials (28) for a meaningful N2pc analysis. All participants reported normal or corrected-to-normal visual acuity.

Experiment 1B: Results

Behavioural results

Preliminary analysis indicated that the shape of the selection feature (square vs. circle) had no effect on accuracy rates, $F < 1$, $BF_{01} = 3.435$, and did not interact with distractor type (letter vs. digit), $F < 1$, $BF_{01} = 2.60$ (adjusting the priors based on the results of Experiment 1A provided stronger support for the null hypothesis in these tests, $BF_{01} = 4.65$ and $BF_{01} = 3.17$, respectively). Therefore, all data were collapsed across this factor. Response accuracy was impaired when the target was followed by a digit distractor relative to when it was followed by a letter distractor (percentage correct: $M = 33.8\%$ vs. $M = 77.0\%$, $t(10) = 11.65$, $p < .001$). When the post-target distractor was a digit, 58.3% of the responses were distractor intrusions (accounting for 88.1% of all errors on these trials). Mean accuracy and intrusion rates are presented in Figure 2B (see Supplementary Figure 2 for individual results).

N2pc components

The average general EEG data loss due to artifacts was 12.5% ($SD = 8.7\%$). Figure 3 (right panels) shows the ERP waveforms triggered by the target frame at electrodes PO7 and PO8 contralateral and ipsilateral to the target on digit distractor trials. The corresponding difference waves obtained by subtracting ipsilateral from contralateral ERPs are shown in Figure 4B (see Supplementary Figure 2 for individual waveforms). Analysis of N2pc onset latencies was based on a 50% average peak amplitude criterion ($M = -1.14 \mu V$). The N2pc component emerged approximately 20 ms earlier on trials with correct responses relative to distractor intrusion trials, $M = 211.9$ ms vs. $M = 232.6$ ms, and this difference was significant, $F_{adjusted}(1,10) = 6.93$, $p = .025$. N2pc mean amplitudes measured in the 200–300 ms time window were significantly different from zero both on trials with correct responses and on distractor intrusions trials, $p < .001$ and $p = .001$, respectively. However, N2pcs were reliably larger on trials where the target

was reported correctly, $t(10) = 4.23, p = .001$.

Discussion of Experiment 1

Experiment 1 yielded two main findings. First, participants' ability to report the identity of the digit target in the RSVP stream was strongly impaired on trials where this target was followed by a digit distractor, and the vast majority of the errors on these trials were distractor intrusions. In fact, such distractor intrusions were more likely than correct reports of the target digit. Second, and most importantly, we demonstrated that such distractor intrusions have a distinct electrophysiological fingerprint. N2pcs elicited by targets that were followed by a distractor digit were significantly delayed on distractor intrusion trials relative to trials with correct responses. These results were nearly identical in Experiments 1A and 1B, except for the fact that the N2pc onset delay on distractor intrusion trials was slightly smaller in Experiment 1B (20 ms, as compared to 30 ms in Experiment 1A). They demonstrate that the speed of attentional engagement processes differed systematically between trials with distractor intrusions and with correct responses. These processes are triggered once the task-relevant selection feature (a specific shape in Experiment 1) has been detected, and the temporal pattern of N2pc components suggest that distractor intrusions are associated with delayed engagement. These results are inconsistent with the claim that distractor intrusions are due to a complete failure to engage focal attention (Botella et al., 2001), and also with the hypothesis that the temporal position of attentional episodes remains essentially identical on trials with correct responses versus distractor intrusions (Vul et al., 2009).

However, before concluding from the N2pc onset latency delay observed on distractor intrusion trials that attentional engagement was delayed on these trials, it is important to consider

alternative interpretations. It is possible that this delay was instead caused by processes that preceded attentional engagement, such as shifts of spatial attention towards the target RSVP stream, and refocusing attention within the attended stream. Experiments 2 and 3 were conducted to investigate these possibilities. In Experiment 1, targets could appear with equal probability and unpredictably in either of the two RSVP streams, and the location of a target was indicated by a shape cue. Once this cue was detected, attention had to be shifted to the relevant stream. Thus, the N2pc onset difference between trials with correct responses and distractor intrusions could reflect differences in the speed with which such attention shifts were triggered. This possibility was investigated in Experiment 3. Furthermore, the selection feature (the shape cue) and the response feature (the identity of the target digit) belonged to different objects in Experiment 1. As a result of this fact, the initial selection of the larger shape cue may have been followed by a recalibration of the attentional focus of attention in order to zoom in on the smaller target object, prior to attentional engagement. A delay of this recalibration process on intrusion trials could in principle have produced the N2pc onset delay observed for these trials. This possibility was assessed in Experiment 2.

Experiment 2

The delayed N2pc observed in Experiment 1 for distractor intrusion trials might not reflect a delayed onset of attentional engagement processes, but instead a slower re-focusing of attention from the shape cue to the digit target object on these trials. If this was the case, such N2pc onset latency differences between correct and distractor intrusion trials should be specific to RSVP tasks where the selection and the response feature are part of different objects, and should not be found when these features belong to the same object. This was tested in Experiment 2 where target digits were defined by their colour. Participants had to report the identity of the first coloured digit that appeared among grey items in one of two RSVP streams. As in Experiment 1, this target digit was followed on the majority of trials by a second digit in the same stream (see Figure 5). The question was whether a sizable proportion of distractor intrusions would be observed on these trials, and, critically, whether these intrusions would again be associated with a delayed N2pc component relative to trials with correct responses. Because the selection and response feature were now part of the same object, no such N2pc onset delay should be found if it was produced by a slower re-focusing of attention.

Another factor manipulated in Experiment 2 was whether the post-target distractor digit was grey or coloured. It is possible that distractor intrusions are more likely to occur for distractors with the task-relevant selection feature. As participants were instructed to report the first coloured digit, coloured post-target distractors matched this task set (“any colour”) while grey distractors did not. If this factor was relevant for distractor intrusion errors, these errors should be more frequent on trials with coloured as compared to grey post-target distractors. In addition, rapid attentional engagement may be more critical for resolving the competition between a target and a subsequent distractor digit when both objects match the selection feature than when the

distractor does not. In this case, any N2pc onset latency difference between trials with correct responses and distractor intrusions may be larger with coloured as compared to grey post-target distractors.

Method

Participants

Participants were 12 (5 women) volunteers ($M_{age} = 28.3$, $SD = 8.6$) who participated for £25. One of them also took part in Experiment 1A. One participant was excluded from analysis because their low accuracy rate on trials where the post-target distractor was a digit (14.6%; 78.8% intrusions) and their rejection rate due to eye movements and eyeblinks (31.0%) left too few trials (30 and 19 for the grey and colour distractor conditions, respectively) for a meaningful ERP analysis. All participants reported normal or corrected-to-normal visual acuity.

Apparatus, stimuli and design

The apparatus, stimuli and design in Experiment 2 were similar to Experiment 1 with the following changes. All items in the RSVP streams were grey, except for the target object and (on some trials) a distractor at the target location in the post-target frame, which were coloured (see Figure 5 for illustration). Outline shapes were not used as selection features, as targets were now defined as the first coloured item encountered in one of the two RSVP streams. These targets were always digits, and participants had to report their numerical value. Target colour was randomly selected in each trial from a set of three colours: blue (CIE colour coordinates: 0.167/.123), green (.306/.615), or orange (.568/.401). All colours were equiluminant (46.6–47.3 cd/m²). The experiment included 800 experimental trials. On 62.5% of these trials (500

trials), the post-target distractor was a digit, whereas the post-target distractor was a letter on the remaining 300 trials. This ratio was chosen to have sufficient numbers of trials for estimating baseline accuracy separately on trials with either grey or coloured post-target letter distractors. Post-target digit or letter distractors were equally likely to be grey (Figure 2A and 2C) or coloured (Figure 2B and 2D). In the latter case, their colour was never identical to the target colour, and was chosen randomly from one of the three remaining colours. In all other aspects, stimulation procedures were identical to Experiment 1.

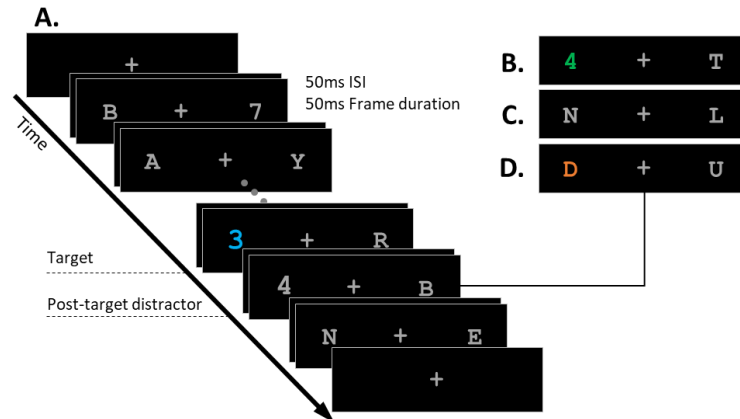


Figure 5. Illustration of the stimulus sequence in Experiment 2. Participants had to report the first coloured digit. The post-target distractor was either a digit or a letter, drawn in grey or colour, as shown in panels A to D.

Behavioural results

Mean accuracy and intrusion rates in Experiment 2 are presented in Figure 2B (see Supplementary Figure 2 for individual results). Accuracy rates were entered as a dependent measure in an ANOVA with post-target distractor type (letter vs. digit) and post-target distractor colour (coloured vs. grey) as within-subject independent variables. As in Experiment 1, participants were more accurate when the post-target distractor was a letter than when it was a

digit, $M = 89.0\%$ vs. $M = 55.8\%$, $F(1,10) = 62.31$, $p < .001$, $\eta_p^2 = .86$. Accuracy was higher on trials where the post-target distractor was grey than when it was coloured, $M = 71.6\%$ vs. $M = 67.0\%$, $F(1,10) = 5.87$, $p = .036$, $\eta_p^2 = .37$, presumably reflecting stronger backward masking on trials where two successive coloured items appeared at the same location. Importantly, the interaction between the two factors was not significant, $F < 1$. When a post-target digit distractor was present, distractor intrusions occurred on 39.5% of trials (accounting for 89.3% of all errors on these trials). No difference in intrusion rates were observed between grey and coloured distractors, $t < 1$ (see Figure 2B).

N2pc components

The average general EEG data loss due to artifacts was 16.1% ($SD = 11.5\%$). Figure 6 shows ERPs triggered by target frames at PO7/8 contralateral and ipsilateral to the target, for trials with a post-target digit distractor. ERPs are presented for trials with correct responses and distractor intrusions, separately for grey and coloured distractors (left versus right panels). The corresponding N2pc difference waves are shown in Figure 4C (see Supplementary Figure 2 for individual waveforms). As in Experiment 1, there was an N2pc onset latency delay on distractor intrusion trials relative to trials with correct responses where distractor intrusions were reported, and this was the case irrespective of whether the post-target distractor was grey or coloured. An N2pc onset latency analysis with a 50% average peak latency criterion ($M = -1.80 \mu V$) was conducted for the factors response (correct vs. distractor intrusion) and post-target distractor colour. This analysis confirmed that the N2pc component emerged reliably earlier for correct trials than for distractor intrusion trials, $M = 180.7$ ms vs. $M = 193.0$ ms, $F_{adjusted}(1,10) = 6.15$, $p = .033$. There was no effect of post-target distractor colour on N2pc latencies, and no interaction between post-target distractor type (letter vs. digit) and colour, both $F_s < 1$. For N2pc mean

amplitude measured in the 200-300 ms post-target time window, there were no reliable main effects of the factors response, $F(1,10) = 1.32$, $p = .28$, $\eta_p^2 = .12$, $BF_{01} = 1.67$, and post-target distractor colour, $F < 1$, $BF_{01} = 3.98$, and there was also no interaction between these two factors, $F < 1$, $BF_{01} = 1.89$. N2pc mean amplitudes were significantly different from zero in all four task conditions, all $ps < .003$.

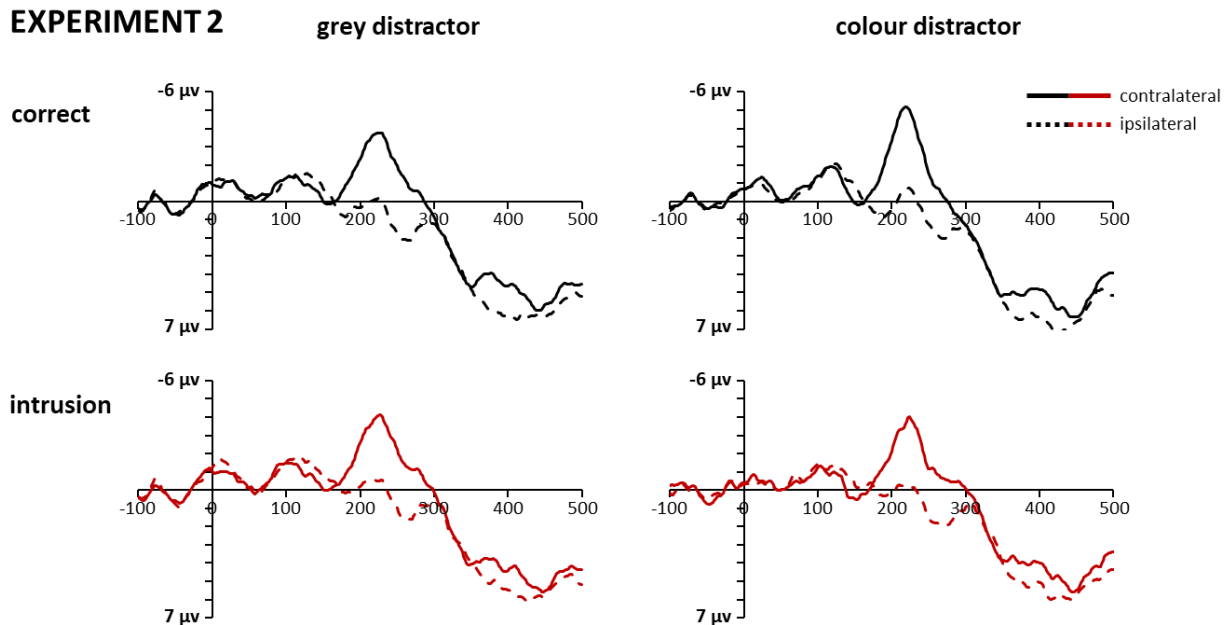


Figure 6. Grand-average event-related potentials (ERPs) waveforms elicited in Experiment 2 by target frames at electrodes PO7/PO8 on contralateral and ipsilateral to the target. ERPs are shown separately for trials with a grey (left panels) or coloured (right panels) post-target digit distractor, and separately for trials with correct responses (top panels) and distractor intrusion trials (bottom panels).

Exploratory comparison between Experiment 1 and Experiment 2

To assess whether changing the selection feature from shape in Experiment 1 to colour in Experiment 2 affected intrusion rates, accuracy rates and N2pc components, the results from these two experiments were compared (after collapsing the data from Experiment 1A and 1B and

collapsing the data Experiment 2 across trials with coloured and grey post-target distractors). In these analyses, we excluded the participant that participated in both experiments from the sample of Experiment 1.

Behavioural results. Distractor intrusion errors were more frequent in Experiment 1 than in Experiment 2, 57.9% vs. 39.5%, $t(31) = 3.014$, $p = .005$. Accuracy was entered as a dependent variable in an ANOVA with experiment as a between-subject independent variable and post-target distractor type (letter vs. digit) as a within-subject independent variable. This analysis revealed a main effect of experiment, as overall accuracy was higher in Experiment 2 than Experiment 1, 72.3% vs. 55.5%, $F(1,31) = 13.04$, $p = .001$, $\eta_p^2 = .30$. As expected, there was also a main effect of post-target distractor type, $F(1,31) = 250.27$, $p < .001$, $\eta_p^2 = .89$, with more errors on trials where this distractor was a digit than when it was a letter, 80.7% vs. 41.5%. The interaction between the two factors did not reach significance, $F(1,31) = 3.548$, $p = .069$, $\eta_p^2 = .10$, though this result was inconclusive, $BF_{01} = 0.89$.

N2pc components. Two between-experiment ANOVAs on N2pc onset latencies and mean amplitudes were conducted with the factors experiment and response (correct vs. distractor intrusion). For the onset latency analysis, the within-group variance of the jackknifed results was first multiplied separately for each group by the relevant $n-1$. This correction adjusts the resulting statistical test (Ulrich & Miller, 2001) while allowing for groups of different size. The analysis obtained a main effect of experiment, $F_{adjusted}(1,31) = 10.89$, $p = .002$, demonstrating that the N2pc emerged earlier in Experiment 2 than in Experiment 1, 187.0 ms vs. 225.5 ms. Even though this delay was numerically larger in Experiment 1 than in Experiment 2 (26.7 ms vs. 13 ms), the interaction between these two factors was not significant, $F_{adjusted}(1,31) = 1.58$, $p = .22$. N2pc amplitudes were larger in Experiment 2 than in Experiment 1, $F(1,31) = 5.028$, $p =$

.032, $\eta_p^2 = .14$. The interaction between experiment and response was significant, $F(1,31) = 5.461$, $p = .026$, $\eta_p^2 = .15$, reflecting the fact that the enhancement of N2pc amplitudes for correct versus distractor intrusion trials was larger in Experiment 1 than in Experiment 2.

Discussion of Experiment 2

Experiment 2 produced two major findings. First, the N2pc onset delay for trials with distractor intrusions relative to trials with correct responses observed in Experiment 1 was replicated, in spite of the fact that the selection feature (colour) and the response feature (numerical value) were now part of the same object. This rules out the possibility that this onset delay was produced by differences in the speed of attentional re-focusing from shape cues to digit targets, and thus supports the alternative hypothesis that distractor intrusion trials were associated with slower attentional engagement processes.

While the comparison between Experiments 1 and 2 was exploratory, it produced two results that are entirely consistent with our temporal variability account. First, target N2pc components emerged earlier and were larger in Experiment 2 relative to Experiment 1. This is likely to reflect faster (and less temporally variable) detection of the target's selection feature, due to (i) the absence of a salient item on the other side which reduced competition between the two streams and (ii) due to the fact that colour is generally more effective in guiding attention than shape (e.g., Wolfe & Horowitz, 2004, 2017), which should result in an earlier onset of attentional engagement processes. Both factors can also account for the second observation that overall accuracy was higher and distractor intrusions were less frequent in Experiment 2 than in Experiment 1. These results are also in line with previous studies showing earlier N2pc onsets for colour-defined targets relative to targets defined in other feature dimensions (Brisson et al.,

2007; Callahan-Flintoft & Wyble, 2017; Töllner, Zehetleitner, Gramann & Müller, 2011) and studies reporting an inverse relationship between the frequency of distractor intrusions and the saliency of target-defining features (Botella, 1992; Chennu, Bowman & Wyble, 2011).

The other notable observation of Experiment 2 was that trials with coloured versus grey post-target distractors did not differ either in terms of N2pc latency differences between correct and distractor intrusion trials, or with respect to the frequency of distractor intrusions. Thus, the speed of attentional engagement and the probability of intrusions was entirely unaffected by whether or not a post-target distractor matched the currently relevant selection feature.

Attentional engagement and distractor intrusions appear to be exclusively determined by processes triggered by the target frame, and not by any additional feature-selective attentional biases for subsequent distractor objects. It should be noted that there was a small but significant reduction of overall response accuracy on trials with a coloured post-target distractor. Since these costs were equal in size regardless of whether the post-target distractor was a digit or a letter, they are likely to be due to stronger low-level backward masking, either due to the inherently higher perceptual saliency of coloured distractors (Ross & Jolicoeur, 1999), or due to fact that coloured distractors matched the attentional task set for colour, thereby enhancing their saliency (Itti & Koch, 2001).

Experiment 3

In contrast to most previous investigations of distractor intrusions, where a single RSVP stream was presented at fixation, Experiments 1 and 2 employed two lateralized RSVP streams. This was necessary to be able to measure lateralized target N2pc components in perceptually balanced bilateral displays. However, because target location was unpredictable in these

experiments, spatial attention could not be allocated to one particular stream in advance. Instead, it had to be shifted to one of the two RSVP streams once the selection feature was detected. Because such attention shifts precede attentional engagement, the N2pc latency delay observed for distractor intrusion trials may not reflect delays in attentional engagement, but instead slower shifts of spatial attention on these trials.

Although N2pc components have been employed to assess attention shifts to target locations (e.g., Ansorge, Horstmann & Worschech, 2010; Jolicœur, Sessa, Dell'Acqua & Robitaille, 2006), previous studies have demonstrated that this component does not reflect spatial attentional orienting as such, but instead a transient process such as attentional engagement that follow attention shifts to task-relevant locations (Kiss et al., 2008; Zivony et al., 2018). For example, Kiss et al. (2008) compared N2pc components elicited by targets in visual search displays in a condition where the location of these targets was cued in advance, so that attention could be shifted to this location prior to the presentation of a search display, and in a condition where target locations were unpredictable, and attention shifts could only be initiated after search display onset. N2pcs were virtually identical in both conditions, demonstrating that this component is not directly linked to attentional orienting. However, any delay in shifting attention to target locations on distractor intrusion trials will also delay subsequent attentional engagement processes. Therefore, the N2pc latency delays found on these trials in Experiments 1 and 2 might reflect not a generic delay of engagement, but instead a knock-on effect of slower spatial attention shifts. If this was the case, our results would not be generalizable to conditions where attention is focused in advance (e.g., as in Vul et al., 2009), as to trial-by-trial variability in the speed of attentional engagement would only affect the perceptual processing of targets and distractors under situations of spatial uncertainty (see also Hogendoorn et al., 2011). In

Experiment 3, we tested this alternative interpretation by making target location fully predictable. Prior to the start of each block, participants were informed that target objects would only appear in one of the two lateral RSVP streams. They were thus able to endogenously shift attention to the task-relevant stream at the start of each trial, and then maintain attention at this location in a sustained fashion. If (i) delays in attentional engagement are not directly associated with distractor intrusions, and if (ii) the N2pc latency delays observed previously for distractor intrusion trials were due to slower attention shifts towards the target object, no such delay should be found in Experiment 3, where no such shifts were required. Alternatively, if there is a direct link between trial-by-trial variability in the speed of attentional engagement and the presence versus absence of distractor intrusions, and if the N2pc delays observed in Experiments 1 and 2 directly reflect slower attentional engagement, analogous delays for distractor intrusion trials should again be observed in Experiment 3.

A sustained focus of attention on one of the two RSVP streams in Experiment 3 may not only affect N2pc onset latencies, but will also result in enhanced sensory responses to all objects within this attended stream. In the EEG, such modulations of visual responses to repetitive stimulation produced by sustained spatial attention give rise to an enhancement of steady-state visual evoked responses (SSVEPs) contralateral to the attended location at the frequency that matches the attended stimulus frequency (e.g., Müller et al., 1998). Because frames within the RSVP streams were presented every 100 ms, allocating spatial attention to one of these streams should result in larger contralateral SSVEPs at the stimulation frequency of 10 Hz.

Method

Participants

Participants were 12 (8 women) volunteers ($M_{age} = 28.17$, $SD = 7.78$) who participated for £25. All reported normal or corrected-to-normal visual acuity. Six of these participants also took part in Experiment 2. Control analyses confirmed that prior participation in Experiment 2 did not affect any of the dependent measures reported here (accuracy rates, intrusion rates, N2pc latencies and mean amplitudes), nor interacted with any independent variable, all $F_s < 1$, all $BF_{01s} > 2$.

Apparatus, stimuli and design

The apparatus, stimuli and design were identical to Experiment 1 (see Figure 1) with the following exceptions. Throughout each block of trials, the target always appeared within the same RSVP stream, either to the right or left of fixation. Participants were informed about this constant target location at the start of each block, which alternated between successive blocks.

Results

Behavioural results

Preliminary analysis indicated that the shape of the target (square vs. circle) had no effect on accuracy rates, $F < 1$, $BF_{01} = 3.38$, and did not interact with distractor type (letter vs. digit), $F < 1$, $BF_{01} = 2.42$, and data were therefore collapsed across this condition (adjusting the priors based on the results of Experiment 1A and 1B provided stronger support for the null hypothesis in these tests, $BF_{01} = 5.41$ and $BF_{01} = 3.43$, respectively). As in Experiments 1 and 2, response accuracy was lower on trials where the target was followed by a digit distractor relative to a letter distractor, $M = 44.2\%$ vs. $M = 82.7\%$, $t(11) = 9.27$, $p < .001$. When post-target distractor was a digit, distractor intrusions occurred on 47.6% of all trials, which accounted for 84.2% of

all errors on these trials. Mean accuracy and intrusion rates are presented in Figure 2C (see Supplementary Figure 2 for individual results).

N2pc components

The average general EEG data loss due to artifacts was 10.8% ($SD = 9.5\%$). Figure 7 shows the ERP waveforms triggered by the target frame at PO7/8 contralateral and ipsilateral to the target, on trials with post-target digit distractors, separately for trials with correct responses and distractor intrusion errors. The corresponding contralateral-ipsilateral difference waveforms are shown in Figure 4D (see Supplementary Figure 2 for individual waveforms). As is evident in these difference waves, a periodic lateralised potential was present in Experiment 3, at a frequency of 10 Hz that corresponds to the frame rate in the RSVP streams. This potential was already present prior to the presentation of the target frame (i.e., during the 100 ms pre-target baseline period), and also during the 500 ms post-target interval (as shown in Figure 4D), where it overlapped with the N2pc components triggered by the target frame. This periodic lateralised component reflects the predicted enhancement of visual processing for each item within the currently attended stream. Notably, no such periodic contralateral response was present in Experiment 1 (Figure 4A and 4B), which was identical to Experiment 3 except that spatial attention was divided across both RSVP streams.

As in Experiments 1 and 2, N2pcs were delayed on distractor intrusion trials relative to trials with correct responses. An analysis of N2pc onset latencies, based on a 50% average peak amplitude criterion ($M = -1.12 \mu V$)³, found an earlier N2pc onset for trials with correct response

³ The overlap of N2pcs with the periodic contralateral SSVEP enhancement associated with focal spatial attention on one of the two RSVP streams might affect N2pc latency comparisons based on a specific amplitude criterion, since these amplitudes also reflect contributions from the early lateralised modulation of sensory-evoked visual

relative to distractor intrusion trials, $M = 177.3$ ms vs. $M = 236.7$ ms, and this difference was significant, $F_{adjusted}(1,11) = 8.43$, $p = .014$. N2pc mean amplitudes measured in the 200–300 ms time window on trials with post-target digit distractors were significantly different from zero both on trials with correct responses and on distractor intrusions trials, both $ps < .001$. Analogous to Experiment 1, N2pc components were larger on trials with correct responses relative to distractor intrusion trials, $F(1,11) = 5.76$, $p = .038$, $\eta_p^2 = .34$.

EXPERIMENT 3

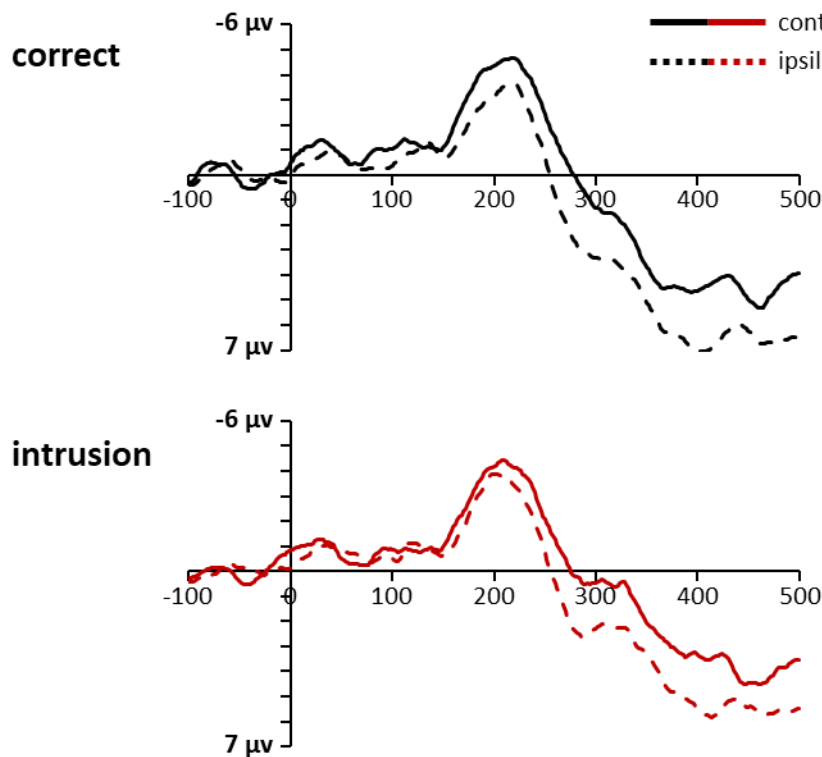


Figure 7. Grand-average event-related potentials (ERPs) waveforms elicited in Experiment 3 on trials with post-target digit distractors by target frames at electrodes PO7/PO8 contralateral and ipsilateral to the target, shown separately for trials with correct responses (top panel) and distractor intrusion trials (lower panel).

activity. However, as RSVP streams were identical on trials with correct responses and on distractor intrusion trials, modulations of early sensory responses by sustained spatial attention should not differ between these trials.

Exploratory comparison between Experiment 1 and Experiment 3

To assess whether the difference between spatially focused attention in Experiment 3 and divided spatial attention in Experiment 1 affected response accuracy and N2pc components, the results from these two experiments were compared.

Behavioural results. Accuracy rates were analysed in an ANOVA with the factors experiment and post-target distractor type (letter vs. digit). There was a main effect of distractor type, $F(1,33) = 310.96$, $p < .001$, $\eta_p^2 = .904$, reflecting more errors on trials with a post-target digit distractor. However, there was no significant main effect of experiment and no interaction between the two factors, $F(1,33) = 2.39$, $p = .13$, $\eta_p^2 = .067$, $BF_{01} = 1.16$ and $F < 1$, $BF_{01} = 2.33$, respectively.

N2pc components. N2pc onset latency was analysed in an ANOVA with the factors response (correct vs. distractor intrusion) and experiment. We applied the same correction as reported above to allow a comparison between unequal groups. Although the mean N2pc latency in Experiment 3 was earlier than in Experiment 1 ($M = 207.0$ ms vs. $M = 227.2$), this difference was not significant, $F_{adjusted}(1,33) = 1.96$, $p = .17$. While the N2pc latency delay on distractor intrusion relative to correct response trials was substantially larger in Experiment 3 relative to Experiment 1 (59 ms versus 27 ms), the interaction between experiment and response only approached significance, $F_{adjusted}(1,33) = 3.77$, $p = .06$. Follow-up analysis indicated that mean N2pc onset latency was earlier in Experiment 3 than Experiment 1 for correct trials, $F_{adjusted}(1,33) = 4.53$, $p = .04$, but not for and intrusion trials, $F_{adjusted} < 1$. For N2pc mean amplitudes, overall N2pc amplitudes did not differ between the two experiments, $F < 1$, $BF_{01} =$

1.73, and there was no interaction between the two factors, $F(1,33) = 1.48$, $p = .23$, $\eta_p^2 = .04$, $BF_{01} = 1.49$.

Discussion of Experiment 3

In Experiment 3, target location was known in advance and remained constant throughout a block of trials, thus obviating the need for transient attention shifts towards target locations. In spite of this fact, the N2pc latency delay for distractor intrusion trials relative to trials with correct responses observed in the first two experiments was again present, suggesting that this latency difference was not primarily due to differences in the speed of attention shifts, and instead reflects trial-by-trial variations in the speed of attentional engagement.

However, two other factors may have contributed to the N2pc latency results observed in Experiment 3. First, it is possible that participants were not able to maintain a constant attentional focus on the target location throughout, and that transient attention shifts to the target were therefore elicited on a minority of trials. If N2pc latency differences reflected differences in the speed of such shifts, one would expect these differences to be considerably smaller relative to Experiment 1, where target locations were unpredictable, and transient attention shifts were therefore required on all trials. This was clearly not the case. If anything, the N2pc latency delay on distractor intrusion trials was numerically larger in Experiment 3. As this comparison was exploratory, we do not want to draw any strong conclusions from this difference between experiments, but it is clearly inconsistent with the possibility that these N2pc latency differences reflect differences in the speed of covert attentional shifts triggered by a selection feature in one of the two RSVP streams. Eye movements are the second factor that may have affected N2pc components in Experiment 3. Because target location was known in advance, participants may

have deviated eye gaze towards this location before the target appeared. This could have distorted the N2pc signal, as the two visual hemifields would no longer be perceptually balanced. However, any such residual drifts in eye position would be problematic only if they differed systematically between correct trials and intrusion trials. An analysis of the HEOG electrodes (see Supplementary materials) revealed no such differences, thus eliminating eye movements as a factor contributing to N2pc latency differences between these two types of trials. Overall, as that neither residual attention shifts nor eye movements towards target location can account for the N2pc results of Experiment 3, they provide further support for our hypothesis that trial-by-trial variability in the speed of attentional engagement is an important factor in explaining distractor intrusions.

The ERP data obtained in Experiment 3 also confirmed that participants followed the instruction to focus covert attention in a sustained fashion on the currently task-relevant RSVP stream. As predicted, ERPs revealed a periodic contralateral attentional response to each successive object within the attended stream at the RSVP stimulation frequency of 10 Hz (see also Müller et al., 1998, for analogous attentional modulations of SSVEPs). Furthermore, while N2pc components emerged about 20 ms earlier in Experiment 3 relative to Experiment 1, and overall response accuracy was numerically higher in Experiment 3 than in Experiment 1, these differences were not statistically significant. Because this could be due to insufficient power to detect such effects in a between-subject design (which was not the main objective of the current study), links between sustained spatial attention, distractor intrusions, and attentional engagement will need to be investigated more systematically in future work.

Finally, the results of Experiment 3 highlight the theoretical importance of treating attentional shifts and attentional engagement as functionally distinct processes. Previously, Zivony and

Lamy (2016; 2018) suggested a “camera” metaphor of attention, where attentional shifts corresponds to the alignment of the lens, and attentional engagement to the shutter-button press. This metaphor is useful to explain performance in situations where attention shifts but attentional engagement is not triggered (Zivony & Lamy, 2018). In the present study, delays in attentional engagement would correspond to pressing the metaphorical shutter-button too late and unintentionally taking a picture of the wrong object (i.e., the distractor). The link between such delays and distractor intrusion errors confirmed in Experiment 3 thus underlines the critical role of attentional engagement in determining perceptual outputs.

General Discussion

The goal of this study was to investigate whether the ability to correctly report target objects in RSVP streams is determined by the speed of attentional engagement processes on individual trials. Participants had to report the numerical value of a target digit that appeared in one of two lateral RSVP streams. In the first two experiments, target location was unpredictable, and targets were defined either by a surrounding shape (Experiment 1) or by their colour (Experiment 2). In Experiment 3, the location of shape-defined targets within one of the two RSVP streams was known and remained constant for an entire block of trials. In all three experiments, when the target was followed by a letter, participants were accurate on approximately 80% of the trials. However, when the target was followed by another digit in the same stream, accuracy in reporting the target was halved. On these trials, participants frequently committed post-target distractor intrusion errors and reported the identity of this distractor digit instead of the target. Distractor intrusions thus reveal a major limitation in temporal selectivity that challenges the widely-held assumption that identifying a single target at high-speed presentation rates is a

highly efficient process (Chun & Potter, 1995; Di Lollo et al., 2005; Jolicoeur & Dell'Acqua, 1998; Taatgen et al., 2009). Critically, in all three experiments, post-target distractor intrusions were associated with significant delays of N2pc components relative to trials with correct responses, strongly suggesting that these intrusions occur when attentional engagement is delayed.

The systematic differences in N2pc onset latency between trials with correct responses and distractor intrusion trials provides clear evidence that the time course of attentional engagement is not constant, but fluctuates considerably across trials. This conclusion contrasts with previous suggestions by Vul et al. (2009) that the temporal position of an attentional window — and the strength of individual object representations within this window — remain essentially constant across trials. According to these authors, distractor intrusions are not linked to any variability in the speed of attentional engagement, but occur when a post-perceptual probabilistic sampling process selects a distractor representation. Their conclusion that trial-by-trial variability in the position of the attentional window is minimal was based on the absence of any predictive relationship between the temporal positions of two items reported on each trial (see above), which is obviously inconsistent with the N2pc onset latency differences between correct and distractor intrusion trials observed here. One way to resolve this discrepancy is to assume that the two reports required in the Vul et al. (2009) study are produced by qualitatively different processes (rapid on-line attentional engagement for the first report, post-perceptual guessing for the second report). This could explain the absence of temporal links between these two reports, as only the first report would be affected by intertrial variability in attentional engagement speed.

The results of our study are also inconsistent with Botella et al.'s suggestion (2001) that distractor intrusions occur on trials where attention fails to be focused on the location of target

objects in RSVP streams. Had this been correct, no evidence for attentional engagement (i.e., no N2pc components) should have been found on distractor intrusion trials. In fact, although N2pcs on these trials emerged later relative to trials with correct reports, they were reliably present in all three experiments, demonstrating that distractor intrusions were not the result of a failure of spatial selection, but were instead associated with a delayed engagement of attention.

It is noteworthy that an additional aspect of the proposal by Vul et al. (2009) is also called into question by the current behavioural results. According to these authors, the probability distribution from which a response is sampled has its maximum for representations of items that coincide with the selection feature, which entails that correct responses will always be selected more frequently than distractor reports. This was not the case in the current experiments. In fact, distractor intrusions were numerically more frequent than correct responses in Experiments 1 and 3. Similarly, this aspect of Vul et al.'s (2009) account may be also incompatible with previous findings from Holcombe and colleagues (Goodbourn & Holcombe, 2015; Goodbourn et al., 2016; Holcombe, Nguyen, & Goodbourn, 2017; Ransley, Goodbourn, Nguyen, Moustafa, & Holcombe, 2018). In these studies, all the distractors shared the target's response dimension, which allowed the calculation of an average positional error relative to the target. Importantly, this average error was consistently positive, which suggests that the centre of the attentional episode was located after target onset. Such a temporal pattern is to be expected, given that attentional engagement is usually triggered as a result of detecting the target, and can be substantially delayed on some trials.

Temporal variability account of attentional engagement

Figure 8 presents a schematic outline of the temporal variability account proposed here, and illustrates how this account can explain the behavioural and electrophysiological results of the

present study. The figure shows how the time course of attentional engagement (fast versus slow) affects the strength of sensory representations in an RSVP task where observers have to report an attribute of a target (the numerical value of a digit) defined by a selection feature (a circle). Feedforward visual processing starts from about 50 ms after stimulus onset (“a” in Figure 8), and generates sensory representations of each stimulus in the RSVP stream within approximately 100-150 ms (Lamme, 2010; Lamme & Roelfsema, 2000). The activation strength of all sensory representations decreases across time as a result of backward masking from subsequent items in the RSVP stream. Attentional engagement is triggered once this process has accumulated sufficient evidence for the presence of the selection feature (at point “b” in Figure 8), and the corresponding sensory representation reaches a threshold criterion (*engagement threshold*; see also Zivony & Lamy, 2018). At this moment, the activation states of all sensory representations at the attended location are transiently amplified via recurrent processing (e.g., Lamme & Roelfsema, 2000), and this is reflected by the emergence of N2pc components. In Figure 8, this transient amplification (indicated by filled areas) is assumed to last approximately 80 ms (Wyble et al., 2009; Wyble et al., 2011). At the end of this attentional amplification period (“c” in Figure 8), the most strongly activated representation at this point is encoded into working memory and thus becomes available for perceptual report.

The critical assumption of this temporal variability account is that the point in time when the engagement threshold is reached varies considerably across individual trials. This variability is due to the fact that pre-attentive perceptual processing is an inherently noisy process (Ashby & Townsend, 1986; Ashby & Lee, 1993), and is also induced by moment-by-moment fluctuations in the generic attentiveness of observers. As a consequence, the selection feature is detected rapidly on some trials and more slowly on others, and the onset of attentional engagement varies

accordingly. When engagement is fast (left panel in Figure 8), the sensory representation of the target is amplified, whereas the representation of the post-target distractor item only becomes available towards the end of the engagement period, and is therefore only weakly activated. When engagement is slow (right panel in Figure 8), the attentional enhancement of the target representation starts later, at a time when this representations is already affected by backward masking, resulting in reduced overall activation levels. In contrast, the representation of the post-target distractor may already be available at the start of the amplification period, and is therefore strongly activated. Crucially, such differences in the onset of attentional engagement across trials result in differences in the perceptual output, as strongly activated sensory representations are more likely to be encoded into working memory. The model illustrated in Figure 8 also assumes that the selection feature is encoded into working memory on all trials. It is then integrated with the representation of the target or the post-target distractor, resulting in a bound percept that forms the basis for perceptual reports (e.g., Chennu et al., 2011; Wyble et al., 2011).

This model can account for all major findings of the present study. First, and most obviously, the observation that N2pc components were delayed on distractor intrusion trials relative to trials with correct perceptual reports in all three experiments is in line with the central hypothesis that distractor intrusions are more likely to occur when attentional engagement processes are delayed. By the time that attentional engagement is triggered (approximately 150-200 ms after the onset of the selection feature), the target representation may be already weakened by feedforward visual processing of the subsequent distractor item. Therefore, even a small delay of 20 ms can bias any competition between these two items in favour of the post-target distractor. In addition, the hypothesis that engagement is triggered once the pre-attentively generated representation of the selection feature reaches a criterion activation threshold predicts that the speed of attentional

engagement is determined by how rapidly a particular selection feature can be detected. It is generally believed that colour is detected faster than shape and other features, and is therefore exceptionally efficient in guiding attention (e.g., Itti & Koch, 2001; Treisman, 2014; Wolfe, 2014; Wolfe & Horowitz, 2004, 2017). In line with this assumption, N2pc components emerged earlier for colour-defined targets in Experiment 2 relative to shape-defined targets in Experiment 1 (see also Callahan-Flintoft & Wyble 2017), indicating that attentional engagement was faster in the former case. The percentage of intrusion errors was significantly lower with colour-defined targets (39.5% ,as compared to 57.9% in Experiment 1), as would be expected if engagement processes were triggered more rapidly.

In all three experiments, accuracy was high on the minority of trials where the target was followed by a letter. This suggests that when the post-target distractor does not match the current response feature, performance is less dependent on the speed of attentional engagement, so that targets are likely to be reported correctly on fast as well as slow engagement trials.⁴ This may suggest a modification of the model outlined in Figure 8. Instead of assuming that only one of the alphanumeric items in the RSVP stream is encoded into working memory (i.e., the item with the highest current activation level), it is possible that both the target and the post-target distractor are encoded, and that the more strongly activated object is then reported. On trials where the post-target distractor is a digit, slow engagement would then again result in frequent distractor intrusion errors. In contrast, when this distractor is a letter, and its identity is therefore

⁴ If correct target reports on trials with post-target letter distractors are independent of engagement speed, target N2pc components will reflect a mixture of fast and slow engagement trials. There should therefore be a tendency for delayed N2pcs on these trials relative to trials with correct responses to targets followed by a digit distractor, where attentional engagement should generally be fast. To test this, we compared target-locked N2pc onset latencies between these two types of trials. In Experiment 1A, the N2pc did indeed emerge significantly later on trials with post-target letter distractors relative to trials with digit distractors, $F_{\text{adjusted}}(1,11) = 4.89$, $p = .049$. However, no such N2pc latency differences were found in Experiments 1A, 2 or 3, all $F_s < 1$.

not part of the response set, error rates should be low even on slow engagement trials, as only one reportable item is present in working memory. The question of how many items are encoded into working memory as a result of attentional engagement in this type of RSVP task will need to be addressed in future studies.

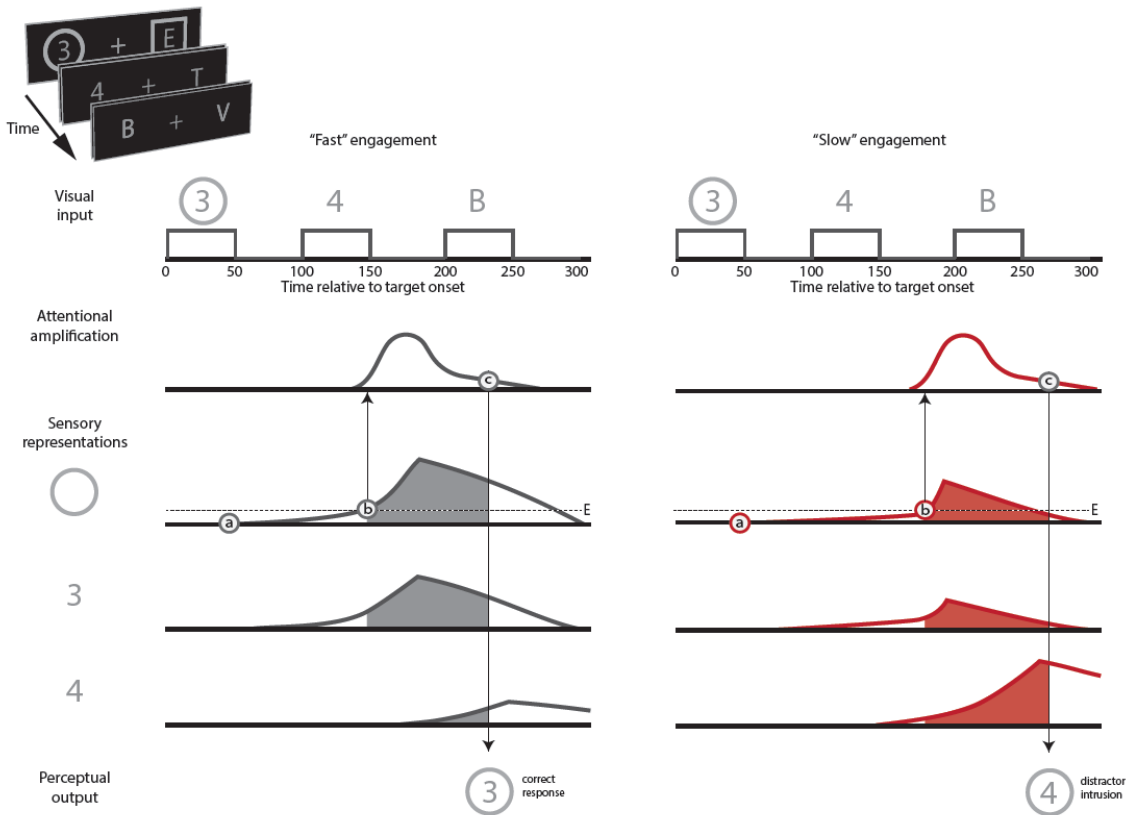


Figure 8. Schematic outline of the temporal variability account. In this example, the selection feature is a circle, the target is “3” and the post-target distractor is “4”. The two panels reflect trials where attentional engagement is fast and the target is correctly reported (left), and trials where engagement is slow and a post-target distractor intrusion occurs (right). Feedforward visual processing starts shortly after stimulus onset (a), resulting in sensory representations of all objects in the RSVP stream. Activation levels initially increase and then decrease due to backward masking. At a specific time point (b), the representation of the selection feature reaches the attentional engagement threshold (E), triggering a transient attentional amplification of all stimulus representations (shaded areas). At the end of this amplification period (c), the most strongly activated digit representation is encoded into working memory and becomes available for report. The point in time where the engagement threshold is reached varies across trials, resulting in a higher probability of post-target distractor intrusion errors when attentional engagement is slow (right panel).

Relationship to related phenomena: Pre-target distractor intrusions, target order reversals, and spatial cuing

The current experiments focused solely on post-target distractor intrusions errors, as digit distractors only appeared after but never immediately before the target object. Many other studies (e.g., Botella et al., 1992; Kikuchi, 1996) found that pre-target distractors can also produce intrusion errors in RSVP streams. According to the temporal variability account proposed here, these intrusions can occur on trials where the perceptual representation of pre-target distractors survive for long enough to be subsequently enhanced by attentional engagement. In line with this account, the frequency of such pre-target distractor intrusions depends on when the selection feature is presented, with more intrusions errors when this feature appears prior to the target object (e.g., Vul et al., 2009). If the probability of intrusion errors depends on the speed of attentional engagement, this should also apply to pre-target intrusions. However, in contrast to post-target distractor intrusions, which are associated with slow attentional engagement and delayed N2pc components, pre-target intrusions should be most likely on trials where engagement is particularly fast, as this will boost the representation of items that immediately precede the target object. Pre-target intrusions should therefore be associated with an earlier N2pc onset relative to trials with correct responses. In addition, because attentional engagement should be faster when target location is known in advance (as suggested by the exploratory comparison between correct responses in Experiments 1 and 3), there should be a paradoxical positive relationship between the predictability of target position and pre-target intrusions, with more such errors occurring when attention is fully focused on one specific RSVP stream. Initial evidence in line with this prediction was provided by Ludowici and Holcombe (2019), who found a negative correlation between the frequency of pre-target

distractor intrusions and the number of RSVP streams in the visual field. This set of predictions will have to be tested in future N2pc experiments where links between attentional engagement speed and pre-target versus post-target intrusion errors is systematically assessed.

The temporal variability account may also serve as a framework to explain a phenomenon that has often been reported in attentional blink experiments. The attentional blink can be observed when two targets in an RSVP stream are separated by approximately 200-500 ms, when observers often fail to identify the second target (Raymond, Shapiro & Arnell, 1992). This effect is much reduced when the two targets appear in immediate succession, without intervening distractors (“lag 1 sparing”; e.g., Visser, Zuvic, Bischof & Di Lollo, 1999). Notably, even when observers report both targets correctly, the second target is often reported before the first (Hommel & Akyürek, 2005). In line with other related attentional prior entry phenomena in the literature (see Spence & Parise, 2010, for review), such order reversals have already been linked to attentional engagement (e.g., Akyürek et al., 2012; Hilkenmeier, Olivers & Scharlau, 2012; Reeves & Sperling, 1986), as pre-cuing the first target reduces their frequency (Olivers, Hilkenmeier & Scharlau, 2011). Our temporal variability account may offer a more specific explanation for these order reversals by interpreting them as the result of differences in the speed of attentional engagement across trials. When attentional engagement is fast, the representation of the first target will be strongly amplified and this target will therefore be perceived first. When attentional engagement is slower, the second target will be activated more strongly, resulting in a perceived order reversal. In other words, the speed of attentional engagement and the resulting relative strength of target activation levels will determine temporal order judgments. If this was correct, order reversals for successively presented targets in attentional blink experiments should be associated with a later N2pc onset relative to trials where both targets are reported in their

correct order. This hypothesis was supported by Callahan-Flintoft and Wyble (2017), who showed that targets which elicited an earlier N2pc were also less likely to be perceived in the wrong order.

Finally, our account highlights the general importance of considering the inherent temporal variability of attentional selection in tasks that involve stimuli that appear in rapid succession. For example, in spatial cuing tasks investigating task-set contingent attentional capture (e.g., Folk & Remington, 1998), target displays are preceded by irrelevant cue displays. Cues that capture attention produce performance benefits when they appear in the same location as the subsequent target. Average differences in the size of these location benefits for different types of cues are often interpreted as evidence for differences in attentional capture (e.g., Becker, Folk & Remington, 2010; Folk & Anderson, 2010; Harris, Becker & Remington, 2015). However, such conclusions are valid if one assumes that the time course of attentional engagement triggered by these cues is relatively constant. Alternatively, different types of cues may induce slower or more temporally variable engagement, which will affect the attentional processing of subsequent targets at cued locations, and thus skew the size of average location benefits. Such temporal variability in attentional engagement may therefore prove to be an important explanatory factor in models of attentional capture and selection mechanisms. This does not only apply to lab-based search tasks, but also to other dynamic situations (e.g., video games) where multiple events appear sequentially.

Conclusion

Distractor intrusions are frequently observed in RSVP tasks, and their presence points to a particular challenge for attentional selectivity under conditions where attention has to be allocated to a specific target object at the right moment in time. Using a combination of

behavioural and electrophysiological measures, we demonstrated that trial-by-trial variations in attentional engagement speed are associated with whether observers can successfully identify the target or report a post-target distractor instead. We found systematic differences in the onset of N2pc components triggered by target frames between trials with correct responses and distractor intrusion trials. The N2pc emerged later on distractor intrusion trials, indicating that intrusions occur when attentional engagement is delayed. These observations demonstrate that subtle variations in the time course of attentional processing can have profound effects on visual perception and performance in situations where multiple objects appear in rapid succession. We propose a new temporal variability account of attentional engagement that can provide a framework for future research into the temporal dynamics of visual object recognition processes.

Acknowledgments

The data for all behavioural analyses and for the N2pc onset latency analysis are posted at <https://doi.org/10.6084/m9.figshare.11837334>. This work was supported by a Newton grant from the British Academy (grant number NIF\R1\180384) to A. Zivony.

Context

The research reported in this article develops and extends previous ideas about the functions of attentional engagement in selective visual processing that were first explored in the PhD work of the first author conducted in Tel Aviv University, and in several publications arising from this work. It also builds the expertise of the second author in employing ERP markers such as the N2pc component to track the time course of attentional selection processes in visual search and visual working memory. This work is part of our ongoing research project on attentional

engagement processes and their role in the adaptive control of different aspects of visual cognition.

References

- Akyürek, E. G., Eshuis, S. A., Nieuwenstein, M. R., Saija, J. D., Başkent, D., & Hommel, B. (2012). Temporal target integration underlies performance at lag 1 in the attentional blink. *Journal of Experimental Psychology: Human Perception and Performance*, 38(6), 1448-1464.
- Ansorge, U., Horstmann, G., & Worschech, F. (2010). Attentional capture by masked colour singletons. *Vision Research*, 50(19), 2015-2027.
- Ashby, F. G., & Lee, W. W. (1993). Perceptual variability as a fundamental axiom of perceptual science. In *Advances in psychology* (Vol. 99, pp. 369-399). North-Holland.
- Ashby, F. G., & Townsend, J. T. (1986). Varieties of perceptual independence. *Psychological review*, 93(2), 154-179.
- Becker, S. I., Folk, C. L., & Remington, R. W. (2010). The role of relational information in contingent capture. *Journal of Experimental Psychology: Human Perception and Performance*, 36(6), 1460-1476.
- Botella, J. (1992). Target-specified and target-categorized conditions in RSVP tasks as reflected by detection time. *Bulletin of the Psychonomic Society*, 30(3), 197-200.
- Botella, J., & Eriksen, C. W. (1992). Filtering versus parallel processing in RSVP tasks. *Perception & Psychophysics*, 51(4), 334-343.
- Botella, J., Arend, I., & Suero, M. (2004). Illusory conjunctions in the time domain and the resulting time-course of the attentional blink. *The Spanish journal of psychology*, 7(1), 63-68.
- Botella, J., Suero, M., & Barriopedro, M. I. (2001). A model of the formation of illusory conjunctions in the time domain. *Journal of Experimental Psychology: Human Perception and Performance*, 27, 1452-1467.

- Brisson, B., Robitaille, N., & Jolicœur, P. (2007). Stimulus intensity affects the latency but not the amplitude of the N2pc. *Neuroreport*, 18(15), 1627-1630.
- Callahan-Flintoft, C., & Wyble, B. (2017). Non-singleton colors are not attended faster than categories, but they are encoded faster: A combined approach of behavior, modeling and ERPs. *Vision research*, 140, 106-119.
- Callahan-Flintoft, C., Chen, H., & Wyble, B. (2018). A hierarchical model of visual processing simulates neural mechanisms underlying reflexive attention. *Journal of Experimental Psychology: General*, 147(9), 1273-1294.
- Chennu, S., Bowman, H., & Wyble, B. (2011). Fortunate Conjunctions Revived: Feature Binding with the 2f-ST2 Model. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 33, No. 33).
- Chun, M. M. (1997). Temporal binding errors are redistributed by the attentional blink. *Perception & Psychophysics*, 59, 1191–1199.
- Chun, M. M., & Potter, M. C. (1995). A two-stage model for multiple target detection in rapid serial visual presentation. *Journal of Experimental Psychology: Human Perception and Performance*, 21, 109–127.
- Di Lollo, V., Kawahara, J., Shahab Ghorashi, S. M., & Enns, J. T. (2005). The attentional blink: Resource depletion or temporary loss of control? *Psychological Research*, 69, 191–200.
- Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological review*, 96(3), 433-458.
- Eimer, M. (1996). The N2pc component as an indicator of attentional selectivity. *Electroencephalography and clinical neurophysiology*, 99(3), 225-234.

- Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2013). G*Power (Version 3.1.7) [Computer software]. Kiel, Germany: University of Kiel.
- Folk, C. L., & Anderson, B. A. (2010). Target-uncertainty effects in attentional capture: Color-singleton set or multiple attentional control settings?. *Psychonomic Bulletin & Review*, 17(3), 421-426.
- Folk, C. L., & Remington, R. (1998). Selectivity in distraction by irrelevant featural singletons: evidence for two forms of attentional capture. *Journal of Experimental Psychology: Human perception and performance*, 24(3), 847-858.
- Folk, C. L., Ester, E. F., & Troemel, K. (2009). How to keep attention from straying: Get engaged! *Psychonomic Bulletin & Review*, 16, 127–132.
- Gathercole, S. E., & Broadbent, D. E. (1984). Combining attributes in specified and categorized target search: Further evidence for strategic differences. *Memory & Cognition*, 12(4), 329-337.
- Goodbourn, P. T., & Holcombe, A. O. (2015). “Pseudoextinction”: Asymmetries in simultaneous attentional selection. *Journal of experimental psychology: human perception and performance*, 41(2), 364-384.
- Goodbourn, P. T., Martini, P., Barnett-Cowan, M., Harris, I. M., Livesey, E. J., & Holcombe, A. O. (2016). Reconsidering temporal selection in the attentional blink. *Psychological Science*, 27(8), 1146-1156.
- Grubert, A., & Eimer, M. (2015). Rapid parallel attentional target selection in single-color and multiple-color visual search. *Journal of Experimental Psychology: Human Perception and Performance*, 41(1), 86-101.

- Grubert, A., Krummenacher, J., & Eimer, M. (2011). Redundancy gains in pop-out visual search are determined by top-down task set: Behavioral and electrophysiological evidence. *Journal of Vision*, 11(14), 1-10.
- Harris, A. M., Becker, S. I., & Remington, R. W. (2015). Capture by colour: Evidence for dimension-specific singleton capture. *Attention, Perception, & Psychophysics*, 77(7), 2305-2321.
- Hilkenmeier, F., Olivers, C. N., & Scharlau, I. (2012). Prior entry and temporal attention: Cueing affects order errors in RSVP. *Journal of Experimental Psychology: Human perception and performance*, 38(1), 180-190.
- Hogendoorn, H., Carlson, T. A., & Verstraten, F. A. (2011). Mapping the route to visual awareness. *Journal of vision*, 11(13), 4-4.
- Holcombe, A. O., Nguyen, E. H., & Goodbourn, P. T. (2017). Implied reading direction and prioritization of letter encoding. *Journal of Experimental Psychology: General*, 146(10), 1420-1437.
- Hommel, B., & Akyürek, E. G. (2005). Lag-1 sparing in the attentional blink: Benefits and costs of integrating two events into a single episode. *The Quarterly Journal of Experimental Psychology Section A*, 58(8), 1415-1433.
- Intraub, H. (1985). Visual dissociation: An illusory conjunction of pictures and forms. *Journal of Experimental Psychology: Human Perception and Performance*, 11(4), 431-442.
- Itti, L., & Koch, C. (2001). Computational modelling of visual attention. *Nature reviews. Neuroscience*, 2(3), 194-203.
- Jolicoeur, P., & Dell'Acqua, R. (1998). The demonstration of short-term consolidation. *Cognitive Psychology*, 36, 138-202.

- Jolicœur, P., Sessa, P., Dell'Acqua, R., & Robitaille, N. (2006). On the control of visual spatial attention: Evidence from human electrophysiology. *Psychological research*, 70(6), 414-424.
- Kikuchi, T. (1996). Detection of Kanji words in a rapid serial visual presentation task. *Journal of Experimental Psychology: Human Perception and Performance*, 22(2), 332-341.
- Kiss, M., Van Velzen, J., & Eimer, M. (2008). The N2pc component and its links to attention shifts and spatially selective visual processing. *Psychophysiology*, 45(2), 240-249.
- Lamme, V. A. (2010). How neuroscience will change our view on consciousness. *Cognitive Neuroscience*, 1(3), 204-220.
- Lamme, V. A., & Roelfsema, P. R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *TRENDS in neuroscience*, 23, 571-579.
- Lins, O. G., Picton, T. W., Berg, P., & Scherg, M. (1993). Ocular artifacts in EEG and event-related potentials I: Scalp topography. *Brain topography*, 6(1), 51-63.
- Luck, S. J. (2014). *An introduction to the event-related potential technique*. MIT press.
- Ludowici, C. J. H., & Holcombe, A. O. (2019). Selection from concurrent RSVP streams: attention shift or buffer read-out?. Poster session presented at the Annual Conference of the Vision Sciences Society, Tampa, FL.
- Miller, J., Patterson, T., & Ulrich, R. (1998). Jackknife-based method for measuring LRP onset latency differences. *Psychophysiology*, 35(1), 99-115.
- Müller, M. M., Picton, T. W., Valdes-Sosa, P., Riera, J., Teder-Sälejärvi, W. A., & Hillyard, S. A. (1998). Effects of spatial selective attention on the steady-state visual evoked potential in the 20–28 Hz range. *Cognitive Brain Research*, 6(4), 249-261.

- Müller, M. M., Picton, T. W., Valdes-Sosa, P., Riera, J., Teder-Sälejärvi, W. A., & Hillyard, S. A. (1998). Effects of spatial selective attention on the steady-state visual evoked potential in the 20–28 Hz range. *Cognitive Brain Research*, 6(4), 249-261.
- Nieuwenstein, M. R., Chun, M. M., van der Lubbe, R. H., & Hooge, I. T. (2005). Delayed attentional engagement in the attentional blink. *Journal of Experimental Psychology: Human Perception and Performance*, 31, 1463–1475.
- Olivers, C. N., Hilkenmeier, F., & Scharlau, I. (2011). Prior entry explains order reversals in the attentional blink. *Attention, Perception, & Psychophysics*, 73(1), 53-67.
- Popple, A. V., & Levi, D. M. (2007). Attentional blinks as errors in temporal binding. *Vision Research*, 47(23), 2973-2981.
- Posner, M. I., & Petersen, S. E. (1990). The attention systems of the human brain. *Annual Review of Neuroscience*, 13, 25–42.
- Potter, M. C., Wyble, B., Haggmann, C. E., & McCourt, E. S. (2014). Detecting meaning in RSVP at 13 ms per picture. *Attention, Perception, & Psychophysics*, 76(2), 270-279.
- Ransley, K., Goodbourn, P. T., Nguyen, E. H., Moustafa, A. A., & Holcombe, A. O. (2018). Reading direction influences lateral biases in letter processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 44(10), 1678-1686.
- Raymond, J. E., Shapiro, K. L., & Arnell, K. M. (1992). Temporary suppression of visual processing in an RSVP task: An attentional blink? *Journal of experimental psychology: Human perception and performance*, 18(3), 849–860.
- Recht, S., Mamassian, P., & de Gardelle, V. (2019). Temporal attention causes systematic biases in visual confidence. *Scientific reports*, 9(1), 1-9.

- Reeves, A., & Sperling, G. (1986). Attention gating in short-term visual memory. *Psychological review*, 93(2), 180-206.
- Ross, N. E., & Jolicoeur, P. (1999). Attentional blink for color. *Journal of Experimental Psychology: Human Perception and Performance*, 25(6), 1483-1494.
- Shih, S. I. (2008). The attention cascade model and attentional blink. *Cognitive Psychology*, 56, 210-236.
- Shih, S. I., & Sperling, G. (2002). Measuring and modeling the trajectory of visual spatial attention. *Psychological review*, 109(2), 260-305.
- Spence, C., & Parise, C. (2010). Prior-entry: A review. *Consciousness and cognition*, 19(1), 364-379.
- Sperling, G., & Weichselgartner, E. (1995). Episodic theory of the dynamics of spatial attention. *Psychological Review*, 102, 503-532.
- Taatgen, N. A., Juvina, I., Schipper, M., Borst, J. P., & Martens, S. (2009). Too much control can hurt: A threaded cognition model of the attentional blink. *Cognitive Psychology*, 59, 1-29.
- Tan, M., & Wyble, B. (2015). Understanding how visual attention locks on to a location: Toward a computational model of the N 2pc component. *Psychophysiology*, 52(2), 199-213.
- Töllner, T., Zehetleitner, M., Gramann, K., & Müller, H. J. (2011). Stimulus saliency modulates pre-attentive processing speed in human visual cortex. *PLoS One*, 6(1), e16276.
- Treisman, A. (2014). The psychological reality of levels of processing. *Levels of processing in human memory*, 301-330.
- Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, 12, 97-136.

- Ulrich, R., & Miller, J. (2001). Using the jackknife-based scoring method for measuring LRP onset effects in factorial designs. *Psychophysiology*, 38(5), 816-827.
- Vierck, E., & Miller, J. (2006). Effects of task factors on selection by color in the rapid serial visual presentation (RSVP) task. *Perception & psychophysics*, 68(8), 1324-1337.
- Visser, T., Merikle, P., & Di Lollo, V. (2005). Priming in the attentional blink: Perception without awareness? *Visual Cognition*, 12(7), 1362-1372.
- Vul, E., Hanus, D., & Kanwisher, N. (2009). Attention as inference: selection is probabilistic; responses are all-or-none samples. *Journal of Experimental Psychology: General*, 138(4), 546-560.
- Vul, E., Nieuwenstein, M., & Kanwisher, N. (2008). Temporal selection is suppressed, delayed, and diffused during the attentional blink. *Psychological Science*, 19(1), 55-61.
- Wagenmakers, E. J., Love, J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., ... & Meerhoff, F. (2018). Bayesian inference for psychology. Part II: Example applications with JASP. *Psychonomic bulletin & review*, 25(1), 58-76.
- Weichselgartner, E., & Sperling, G. (1987). Dynamics of automatic and controlled visual attention. *Science*, 238(4828), 778-780.
- Wolfe, J. M. (2014). Approaches to visual search: Feature integration theory and guided search. *The Oxford handbook of attention*, 11, 35-44.
- Wolfe, J. M., & Horowitz, T. S. (2004). What attributes guide the deployment of visual attention and how do they do it?. *Nature reviews neuroscience*, 5(6), 495-501.
- Wolfe, J. M., & Horowitz, T. S. (2017). Five factors that guide attention in visual search. *Nature Human Behaviour*, 1(3), 0058.

- Woodman, G. F., & Luck, S. J. (1999). Electrophysiological measurement of rapid shifts of attention during visual search. *Nature*, 400(6747), 867-869.
- Wyble, B., Bowman, H., & Nieuwenstein, M. (2009). The attentional blink provides episodic distinctiveness: sparing at a cost. *Journal of experimental psychology: Human perception and performance*, 35(3), 787-807.
- Wyble, B., Potter, M. C., Bowman, H., & Nieuwenstein, M. (2011). Attentional episodes in visual perception. *Journal of Experimental Psychology: General*, 140(3), 488-505.
- Zivony, A., & Lamy, D. (2016). Attentional capture and engagement during the attentional blink: A “camera” metaphor of attention. *Journal of Experimental Psychology: Human Perception and Performance*, 42(11), 1886–1902.
- Zivony, A., & Lamy, D. (2018). Contingent attentional engagement: stimulus- and goal-driven capture have qualitatively different consequences. *Psychological Science*, 29, 1930-1941.
- Zivony, A., Allon, A. S., Luria, R. & Lamy, D. (2018). Dissociating between the N2pc and attentional shifting: an attentional blink study. *Neuropsychologia*, 121, 153-163.

Supplementary Materials

HEOG analysis

To ensure that small eye movement did not create any consistent N2pc differences between correct trials and intrusion trials, we analyzed the remaining HEOG data following rejection of large eye movements. Similar to the main analysis, we included only trials where the target was followed by a digit. For each participant, we calculated the average difference wave between the two HEOG electrodes when they were ipsilateral versus contralateral to the target's visual field, separately for correct trials and intrusion trials. In the resulting difference waves, positive deflections reflect an deviation of eye gaze towards the target. Next, we applied a 10-hz low-pass filter. We then compared the waveform for correct trials and intrusion trials on two criteria: (i) mean amplitude in the N2pc time window (200-300 ms), and (ii) mean onset latency. For the latency analysis, we applied the jackknife procedure and corrected the statistical test according to the formulas provided by Miller, Patterson, and Ulrich (1998; Ulrich & Miller, 2001). The onset latency criteria was defined as the point where the difference waveform reached 50% of the average peak amplitude across trials with correct responses and distractor intrusion trials. Note that the application of this filter on segmented data creates edge artifacts, but these do not affect our analysis that focuses on the N2pc time window. Supplementary Figure 1 reflect the HEOG difference wave for Experiments 1-3.

Experiment 1A

As can be seen), HEOG deflections reflecting small eye gaze deviation towards the target location emerged only after the N2pc time window. The mean onset latency was 373 ms for correct trials and 400 ms for intrusion trials. The latency difference between these conditions was not significant, $t_{adjusted} < 1$. The average amplitude of the HEOG difference wave during the N2pc time window was 0.03 μ V for correct trials and -0.02 μ V for intrusion trials (which corresponds to an average eye deviation of less than 0.01°), and this difference was non-significant, $t < 1$, $BF_{01} = 3.36$,

Experiment 1B

HEOG deflections again only emerged after the N2pc time window (Supplementary Figure 1, upper right). The mean onset latency was 323 ms for correct trials and 344 ms for intrusion trials. The difference between these conditions was not significant, $t_{adjusted}(11) = 1.37$, $p = .20$. The average amplitude of the HEOG difference wave during the N2pc time window was 0.28 μV for correct trials and 0.038 μV for intrusion trials, which corresponds to an average eye gaze shift of less than 0.02° . While the difference between these mean amplitudes was statistically significant, $t(11) = 2.25$, $p = .0485$, the influence of these small eye gaze deviations on the waveforms recorded in the PO7 and PO8 electrodes is negligible (Lins, Picton, Berg & Scherg, 1993) and cannot explain any ERP differences between these conditions.

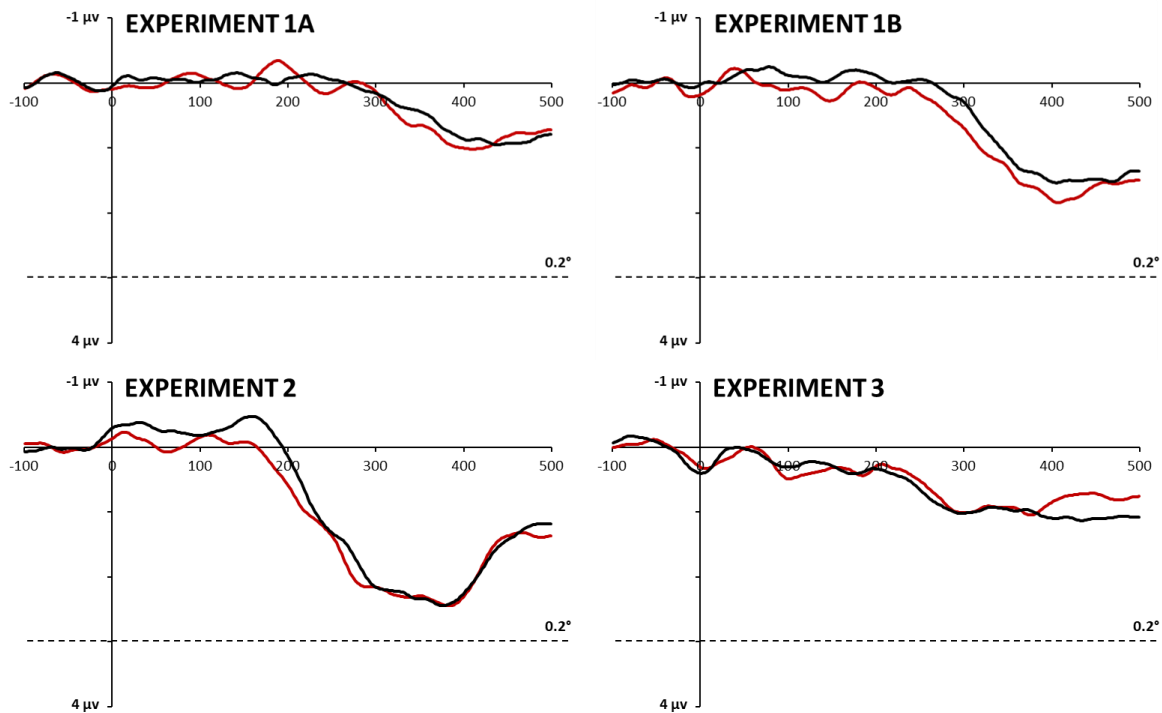
Experiment 2

As can be seen from Supplementary Figure 1 (lower left), a deflection in the difference wave, reflecting microsaccadic eye movements, now emerged already during the N2pc time window. The mean onset latency for this deflection was 213 ms for correct trials and 227 ms for intrusion trials, but this difference was not significant, $t_{adjusted}(11) = 1.45$, $p = 1.45$. The average amplitude of the HEOG difference wave during the N2pc time window was 1.48 μV for correct trials and 1.28 μV for intrusion trials. Although this was larger than in Experiment 1, it still reflects an average eye gaze deviation of less than 0.1° for both trial types, Lins et al., 1993). Most importantly for the current purposes, the HEOG amplitude difference between correct and incorrect trials was not significant, $t(10) = 1.09$, $p = 0.3$, $BF_{01} = 2.07$, and thus cannot account for any N2pc differences between these trials.

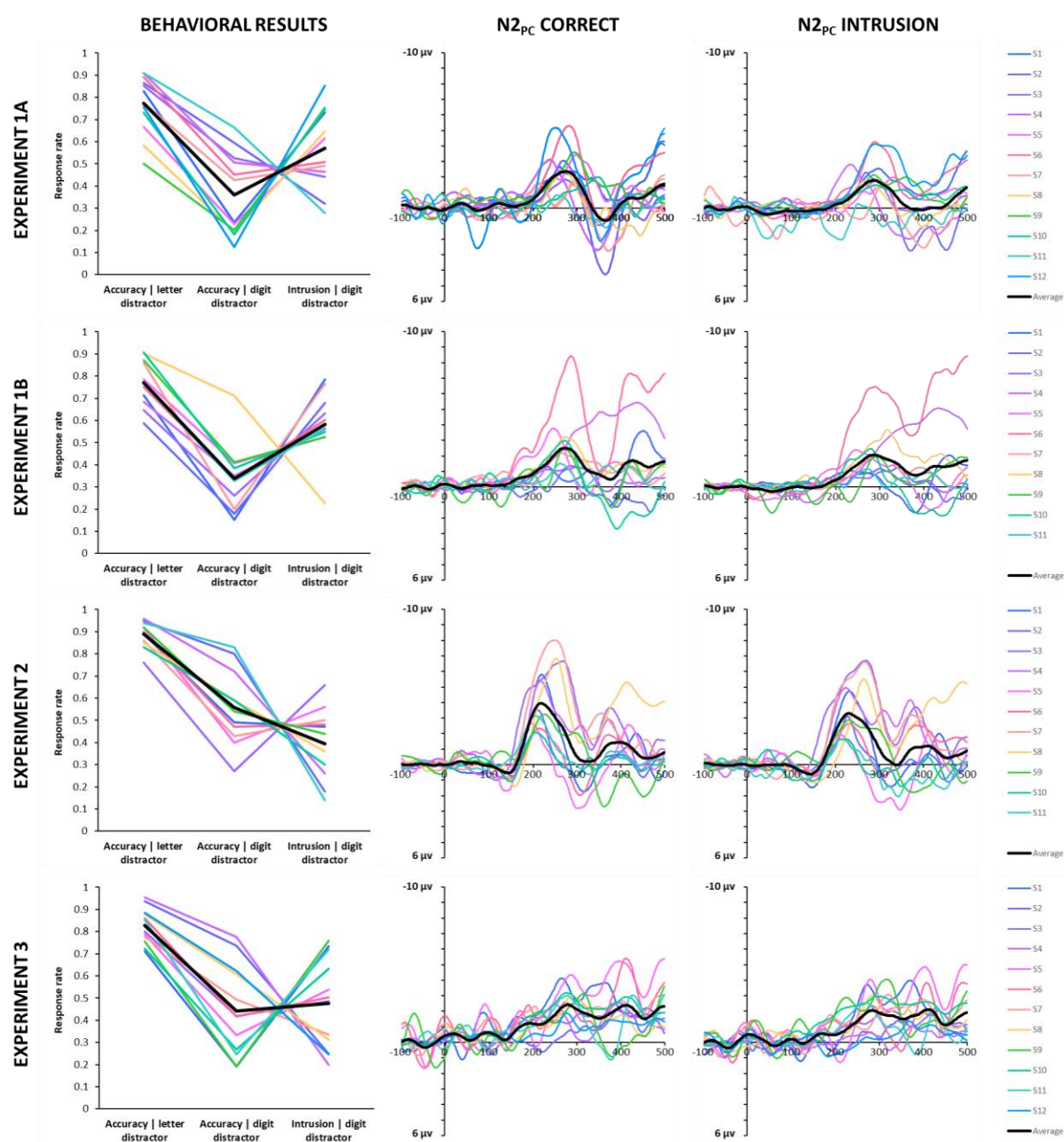
Experiment 3

As shown in Supplementary Figure 1 (lower right), small eye gaze deviations towards the target emerged late during the the N2pc time window, with average onset latencies of 277 ms for correct trials and 268 ms for intrusion trials. This latency difference was not significant, $t_{adjusted} < 1$. The average amplitude of the HEOG difference wave during the N2pc time window was 0.59 μV for correct trials and 0.68 μV for intrusion trials. The difference in mean amplitudes was not significant, $t < 1$, $BF_{01} = 3.12$.

Overall, these HEOG analysis showed that differences in eye gaze deviations towards targets between different trial types cannot account for the pattern of N2pc results found in our study. In Experiment 1A and 1B, consistent eye movements towards the target emerged only after the N2pc time window. In all three experiments, there was no difference between correct and intrusion trials in either the onset latency or mean amplitude of HEOG difference waves during the N2pc time window. As can be seen from Supplementary Figure 1, none of the HEOG difference waveforms reached 3 μ V (dotted line). This suggest that after artifact rejection, average eye gaze deviations remained below 0.2° in all experiments.



Supplementary Figure 1. HEOG difference waves for experiments 1-3, calculated as the difference between HEOG electrodes contralateral and ipsilateral to the visual field of the target, shown separately for trials with correct responses (black lines) and intrusion errors (red lines). The dashed lines represents a HEOG deflection that corresponds to an average eye gaze deviation of 0.2°.



Supplementary Figure 2. Behavioural and N2pc results for individual participants. Thin coloured lines reflect individual data and the thicker black lines reflect averaged. Behavioural results (leftmost column) show response rates (accuracy versus intrusion rates) as a function of post-target distractor type (letter versus digit). The N2pc results are shown only for trials where the target was followed by a category-matching distractor, separately for correct trials (middle column) and intrusion trials (rightmost column).